

# FT-IR Measurement of PSG and BPSG Films

## Part II: Partial Least-Squares Approach

Fred Walder, Ross Boyle, Thermo Fisher Scientific, Madison, WI, USA

### Introduction

The move from LSI to VLSI devices has placed more critical requirements on semiconductor process conditions. Narrow line widths and the accompanying size reduction of all device topographic features make uniformity and reduced process temperatures important. In interlayer dielectric films, the use of boron and phosphorus dopants in silicate glass has become prevalent. The addition of these dopants lowers the temperature required to soften, or flow, the glass layer, minimizing diffusion during reflow. Consistent and uniform deposition of the chemical vapor deposition (CVD) glass assures higher device yield.

The optimization of reactor performance and the need for routine film evaluation make a quick determination of these dopant levels desirable. Though several methods have been utilized, the advantage of Fourier Transform Infrared (FT-IR) measurement is that boron phosphorus and glass thickness can be simultaneously determined in a relatively fast (one minute), non-contact procedure.

Classical approaches to this problem have yielded acceptable results in some cases, but as shown in Part I<sup>1</sup> of this report, considerable difficulties are associated with predicting the concentration of boron and phosphorus over large ranges. These difficulties arise as a result of the changes in spectroscopic features associated with changing dopant levels. Specifically, in the analysis of BPSG films, there are three components of interest. But greater than three spectral components are needed to model and predict concentrations of unknown films.

Significant improvement in determining unknown film concentrations is possible using the Partial Least Squares (PLS) algorithm. The basic principles of the technique, as well as several of the important procedures, are described as follows.

PLS is a statistical quantitative method of analysis. Its major applications arise in situations where mixtures of materials are complex, for example:

- the components in mixture form take on unique qualities (spectral features) and are a function of the presence and concentrations of other components of the mixture; or
- direct application of Beer's Law (based methods) does not provide sufficiently accurate prediction values for multicomponent systems.

The user has some choices about spectral activity; however, the PLS program automatically chooses the regions or spectral features that provide the greatest ability to predict concentrations. This is in contrast to the K-matrix method in which proper band choices are made in advance of calibrating or using the program and are mandatory for the success of the method.

The PLS technique reduces a set of calibration spectra into a smaller number of factors. The factors can be presented as a spectrum, using intensity (absorbance) versus wavenumber (frequency). The calibration spectra may be represented by the following equation:

$$A = TB + e \quad (\text{Eq. 1})$$

where:

- A = the measured absorbance spectrum
- B = loading spectra, or factors
- T = intensities in the new coordinate system
- e = random errors or residuals not fit by the new model

The relationship between Equation 1 in this report and Equation 5 ( $A=KC$ ) in Part I<sup>1</sup> is noteworthy. In the Beer's Law based K-matrix approach, the observed absorbance spectrum is defined as a product of the individual component spectra and a concentration matrix. In the above equation, the observed absorbance spectrum is shown to be composed of the product of loading spectra, or factors, and an intensity matrix.

A similar equation may be considered for concentrations:

$$C = UD + e \quad (\text{Eq. 2})$$

where:

- C = the concentration matrix
- D = a loading matrix for concentrations
- U = intensities in the new coordinate system
- e = random errors or residuals not fit by the new model

Understanding the B-matrix, or loading spectra, is the point of interest. The PLS algorithm defines a set of loading spectra as linear combinations of spectral features from the measured set of calibration spectra. These factors are ranked by their ability to predict the concentrations. The estimation of each factor is carried out in such a way as to minimize the residuals in concentration and in spectral information. The contribution of the factor to spectral and concentration data is then removed for that given factor, and the calculation is repeated for the residuals, and so on. Qualitatively stated, the loading spectra are spectral constructs designed to enable the prediction of sample concentrations. They are statistically evaluated on their ability to minimize residuals in spectral and concentration

### Key Words

- BPSG
- Dielectric Films
- ECO
- FT-IR
- PLS
- PSG
- Semiconductor Metrology

information (term  $e$  in Equations 1 and 2). Spectral information is used in the construction; however, it is important to remember the factors represent statistical solutions based on predictive ability.

Although PLS can be “taught” to deal very successfully with band variability (for example, as deposited films) by using the correct references, standardizing the films with an anneal reduces the number of references required. Thermal cycling the films at their reflow temperature densifies the glass, relieves the stress in the film and drives all species of interest into the optimum oxidation state for infrared detection. The last benefit is evident when viewing the spectrum because almost no evidence of a phosphorus absorbance can be seen before anneal, and the main Si-O band is shifted and broadened.

## Experimental

In order to more fully develop the above concepts, data from an actual sample set of plasma enhanced chemical vapor deposition (PECVD) films deposited on high-resistivity substrates and densified will be considered. Twenty-two sample spectra were analyzed: half were run as calibration samples and half as validation samples. The purpose of the validation set is to verify the performance of the method on spectra not used as calibration data, and to optimize the use of the loading spectra. The samples varied in glass thickness from 3800 Å to 5400 Å; boron wt. % varied from 1% to 4%; and phosphorus varied from 3.3% to 9%. Energy dispersive X-ray analysis was used to determine the actual dopant concentrations.

## Results and Discussion

The first step in the PLS method is to obtain data for the calibration and validation. Preferably this is done on identical samples that are analyzed later by independent techniques, such as wet chemical analysis or X-ray analysis, to determine actual concentration. In the case of the data presented, half wafers were used for IR calibration with the sister halves having been analyzed by an independent technique.

The PLS method is a statistical model in which the computer program develops the analysis based on optimal spectral areas, so a larger number of references is preferred. Though the PLS method will function mathematically with smaller numbers of references, we recommend a minimum of 11 arranged in a manner surrounding the concentration of interest. One possible pattern is illustrated in Figure 1,

showing a heavier concentration of calibration standards in the target representing expected concentrations for deposited films. Calibration standards should be samples perfectly representative of the films to be analyzed. Generally, if larger windows or target areas of

dopant concentration or glass thickness are expected, more calibration design points are required.

The importance of the proper calibration set to statistical methods of analysis such as PLS can not be overstated. In band assignment techniques such as K-matrix, locations have been assigned in the spectrum to relate to concentration information. Though this is conceptually more straightforward, it suffers from two problems for the analysis of BPSG films: the presence of boron and phosphorus affect each other in the glass matrix, thereby changing each other's band shapes and location, and the shape and position of the glass band changes with dopant concentration.

In the PLS approach, no band location decisions are made in advance. A general area of the spectrum presenting high activity may be selected for the entire analysis (the 1600 to 700  $\text{cm}^{-1}$  area is routinely recommended), but no specific band assignments are called for or are, in fact, useful. Because PLS “assigns” the spectral features correlating to concentration, and because more spectral components are present than the number of known pure components used in a band height calculation, a more complete set of references is necessary.

Another way to look at this is that the decomposition of the absorbance spectrum into individual concentrations takes place in both methods, but in PLS, there are considerably more “B” terms than the corresponding “K” terms in the K-matrix approach. This method is useful because the interaction of components in the mixture requires consideration of more than just the pure component or calculated pure component terms.

Actual band assignments are not called for in the setup of the PLS method, but it is often useful to examine spectra of various concentrations to determine areas of spectral activity: that is, areas of the calibration spectra shown to change from sample to sample. Several alternate methods are available for this. First, one may look at actual data. Figure 2 shows three BPSG curves from different concentration samples used in the analysis.

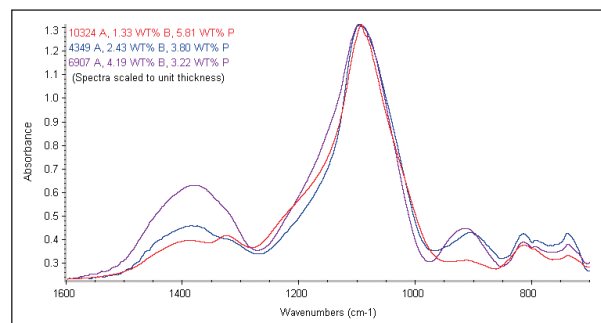


Figure 2: Three different BPSG films with varying concentrations of Boron and Phosphorus

The area of spectral activity for BPSG films is confined to the 1600 to 700  $\text{cm}^{-1}$  range. When lower than 700  $\text{cm}^{-1}$ , the signal-to-noise is poor and the spectrum from the silicon substrate can interfere with the calibration. When higher than 1600  $\text{cm}^{-1}$ , there is less relevant information and water vapor tends to interfere with the calculation unless the system is maintained under perfect purge conditions.

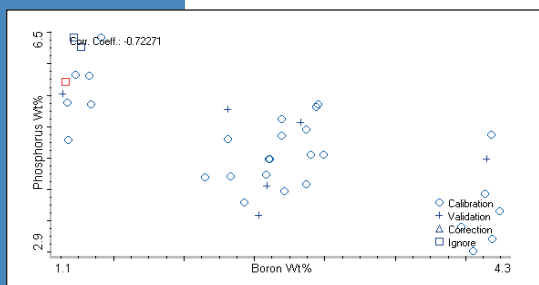


Figure 1: Chart showing concentration design points for proper calibration of the PLS method as applied to BPSG films

Additional tools of the PLS method are variance and correlation spectra, which are used to determine areas of spectral activity. Large changes are present in the upper curve region; however, the variance and correlation display present the areas of most spectral change in the reference set, with and without correlation to the changing concentration values. In concert, these can also be useful in defining regions for analysis.

Correlation spectra can either be generated for single components or can represent the correlation to the changes of all defined concentrations in the range of interest. An example of the multiple-component correlation spectrum is presented in Figure 3. In comparing the raw variance display, multiple- and single-component correlation spectra can provide useful qualitative information on chemical structure changes with changing concentrations.

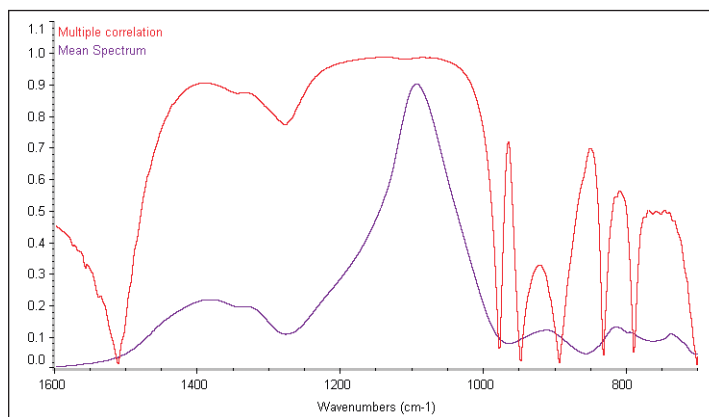


Figure 3: Plot showing regions of high correlation of spectral differences to changes in concentration of all components

Several aspects of the correlation display presented in Figure 3 are intuitive from examination of the actual data, whereas several are surprising. The large variation in the 1400  $\text{cm}^{-1}$  area obviously comes about as a result of the correlation of the change of that area and the boron concentration. This not only reinforces the band often assigned for the B-O stretching vibration, but also supports the fact that even with less sophisticated software methods, the analysis of boron is possible. An examination of the correlation spectrum for boron alone, not presented here, readily confirms this.

The traditional phosphorus area of analysis, at about 1310  $\text{cm}^{-1}$ , is actually seen as a rather minor inflection. Its small size when compared with the changes taking place in the reference spectra gives some indication of the difficulties encountered when trying to determine phosphorus by use of this band alone.

The apex of the Si-O band, traditionally analyzed in the 1050-1100  $\text{cm}^{-1}$  area, is actually a negative inflection in the correlation display. In addition, the shoulders of the Si-O band in the 1200 and 1000  $\text{cm}^{-1}$  areas are shown to be changing significantly, with a positive correlation to concentration changes. This information gives the first strong sign (aside from a failure to adequately perform the analysis by other methods) that simple band assignment does not present the complete story of the possible changes in these spectra.

Though examining individual correlation spectra is possible, the next presented step is to allow the PLS program to generate the loading spectra (B term in Equation 1). The program generates the number of factors necessary to minimize the error in predicting the concentration of the validation samples. In this implementation of the PLS technique, different factors may be used to predict the concentration of different individual components.

In the analysis, a set of nine loading spectra was generated. Table 1 lists the loading factor, the cumulative percent of spectral information accounted for by using all factors up to that number, and the cumulative amount of concentration information for the individual components accounted for by all factors up to that number. Stated in another way, term e in Equation 1 would be 100 minus the percent listed in the % Spectra column for the utilization of all the spectral data in the validation set (term A in Equation 1). Term e in Equation 2 would be 100 minus the values in the % thickness of the boron and phosphorus columns, respectively, in attempting to fit the C-matrix information of Equation 2. It is the consideration of both these matrices that characterizes the unique qualities of the PLS approach, and contrasts it with other statistical methods.

FACTOR #	% SPECTRA	% THICKNESS	% BORON	% PHOS	THICKNESS LOADING	BORON LOADING	PHOS LOADING
1	28.90	51.47	53.69	1.61	0.695	0.704	0.142
2	53.65	66.14	68.78	66.25	0.332	-0.327	0.884
3	70.18	75.53	85.80	66.75	0.635	0.755	0.157
4	81.78	82.24	92.25	85.19	0.464	0.328	0.822
5	89.15	87.90	92.69	87.80	0.828	0.165	0.535
6	91.90	<b>95.18</b>	93.27	92.48	0.742	0.189	0.642
7	95.58	95.19	95.80	92.51	-0.073	-0.989	0.127
8	96.95	95.45	<b>99.08</b>	97.29	-0.188	-0.496	0.847
9	98.57	97.98	99.57	<b>99.38</b>	0.828	0.165	-0.535

Table 1: Data showing the amount of spectral and concentration information used with the addition of each successive factor, and weightings for each component to each factor

In addition, the significance of each component fit to each factor is listed. Positive loadings occurring with large changes in predictive ability indicate the significance of a particular factor or loading spectrum to the prediction of concentrations of that component. This will be examined further as the individual factor loadings are presented in the spectral domain.

Different factors may be selected to predict each component. In this data set, the boldface numbers indicate that six factors were used for prediction of thickness, eight for boron and nine for phosphorus. These numbers are determined by the error magnitude in predicting the concentrations of the validation samples. Generally, the error magnitude goes through a minimum at a particular number of factors. The error magnitude is determined by the ability to predict the concentration values of validation samples, not of the calibration samples, so the minimum value indicates the number of factors that would be optimal for predicting unknown sample concentrations. Here the value of having a validation set can be seen. If only the concentration of the calibration samples were evaluated, the largest number of factors would always be optimal because of the non-random spectral variations associated with the calibration spectra (noise, purge differences, etc.).

Qualitatively, one could state that the interactions and distortions requiring the use of additional factors beyond the simple three-for-three components are less complex for thickness prediction, and more complex for phosphorous. This is not unreasonable based on the degree of band overlap and the relative success of older quantitative methods for predicting BPSG.

Term B from Equation 1, or the individual factors or loading spectra, can be examined in the form of loading intensity versus wavenumber. This is of interest to understand the significance of certain factors to prediction. The use of factors representing mixtures of all three components of interest makes excessive qualitative interpretation of the individual factors more difficult; however, there is often some information to be gathered during method development.

An example is presented in Figure 4, where the first factor is shown. It represents a level accounting for 28.90% of spectral information (A) used and just over 50% of the boron and thickness information from the concentration matrix (C). It offers little to the analysis of phosphorus. The high-positive, component-loading values also verify its significance for boron and thickness prediction. The ordinate for the factor plots presented is loading intensity in the new coordinate system. Though they all vary to some extent, they are of similar magnitude and contain approximately equal amounts of positive and negative information.

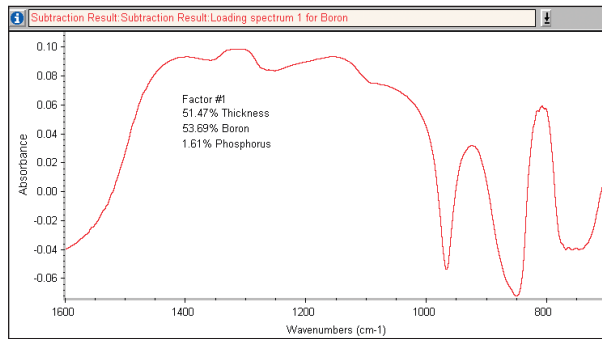


Figure 4: Plot showing the first factor, or loading spectrum

Other qualitative judgements can be made by examining Figure 4. The 1390 and 925  $\text{cm}^{-1}$  areas of the calibration spectra, which change with boron concentration, match two of the strongest three features in the first factor. The third strong feature is in the 1200  $\text{cm}^{-1}$  area corresponding to the change in the shape of the Si-O band with thickness of the glass film. Though accounting for just less than 30% of the spectral information, this first factor is comprised of significant information correlating to the changes that boron and thickness cause in the spectrum.

Table 1 shows that Factor 2 is significant for the prediction of phosphorus, with both a large increase in percent of concentration predicted and a large positive loading for phosphorus. A similar conclusion could be drawn for boron and Factor 3. Graphical presentations of both these factors are found in Figures 5 and 6.

Factor 2 shows a large positive feature in the area of the P=O band and on the shoulders of the glass band. It also shows negative features in the locations typically associated with the boron bands.

Following the addition of Factor 3, over 70% of the available spectral information (A) in the validation data set has been accounted for. Noisy features will be more prevalent in later factors as a result of the smaller amount of spectral information left to fit for these factors. It is,

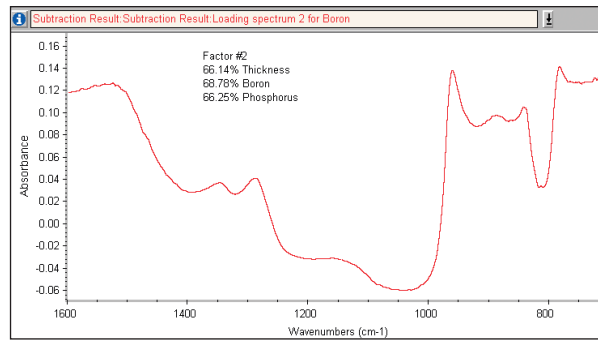


Figure 5: Plot showing the second factor, or loading spectrum

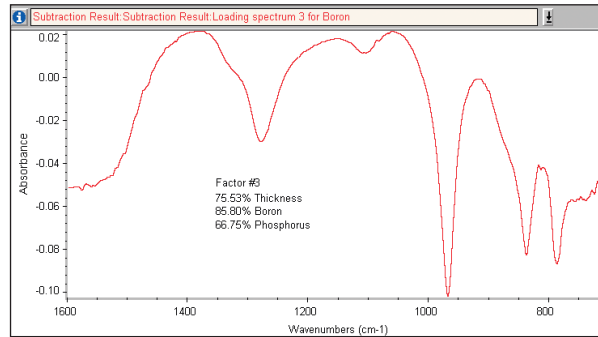


Figure 6: Plot showing the third factor, or loading spectrum

however, their representing an enhancement in the ability to predict the concentration matrix (or minimize term e in Equation 2) that is significant.

The final factor to be examined is Factor 8 in Figure 7. Factor 8 is not used for thickness determination, but does represent a large improvement in the prediction of phosphorus concentrations. A large positive loading value for phosphorus is present for that factor, making the phosphorus analysis much more reasonable; both boron and phosphorus prediction errors decrease by over 60% with addition of this factor. A derivative shape in the plot of Factor 8 shows a positive maximum in the area above 1300  $\text{cm}^{-1}$  where the phosphorus has its strong IR absorption. The plot also shows the effect that the addition of phosphorus could have on both the glass band maximum and the shape of the leading edge of the glass band.

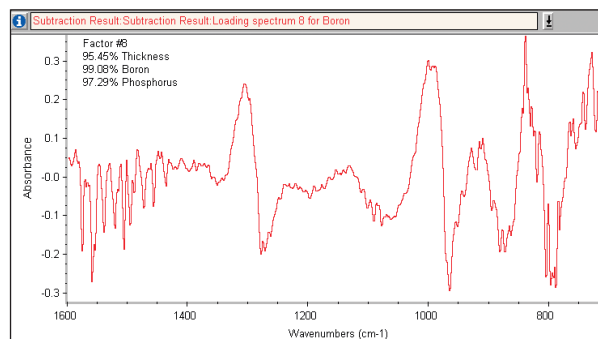


Figure 7: Plot showing the eighth factor, or loading spectrum

As mentioned, larger factor numbers tend to represent less spectral information and are generally more significant in the prediction of the concentration matrix. Only 1.37% of the spectral information is found in Factor 8 because of

its “noisy” appearance. This significantly improves the ability to predict boron and phosphorus concentrations. The ability of a factor to account for concentration information (C) versus spectral information (A) is called the weighting. As the amount of concentration information increases relative to spectral information, the weighting increases. Each of the factors in the analysis has a cumulative weighting for its contribution to concentration prediction. For Factors 1, 2, 3, 4 and 8, the weighting values are 0.118, 0.127, 0.075, 0.102 and 0.178, respectively. Interestingly, Factor 8 represents the maximum weighting value for the analysis. The minimum was 0.056 and belonged to Factor 7 (not shown). Further data analysis is available in the PLS package to examine the contributions of individual calibration spectra to each factor and the weighting calculation, but this will not be presented here.

Though the available diagnostic and analytical tools are beneficial, it is important to examine whether PLS can improve the predictive ability for BPSG samples and to examine the results. With the described data set, the average prediction errors can be summarized as seen in Table 2.

#### METHOD AVERAGE PREDICTION ERROR

	Thickness	Boron	Phosphorus
K-matrix	290 Å	0.4 wt. %	1.3 wt. %
PLS	41 Å	< 0.1 wt. %	0.17 wt. %

Table 2: Data showing the improvement in predictive ability that the PLS method lends to the analysis of PECVD BPSG films

The PLS method increases the ability to predict glass thickness by a factor of five over more conventional band-height-assignment methods. It also improves the boron ability by a factor of four, and finally makes the prediction of phosphorus possible, as FT-IR readings of P in BPSG have been quite difficult to correlate with previous methodologies.

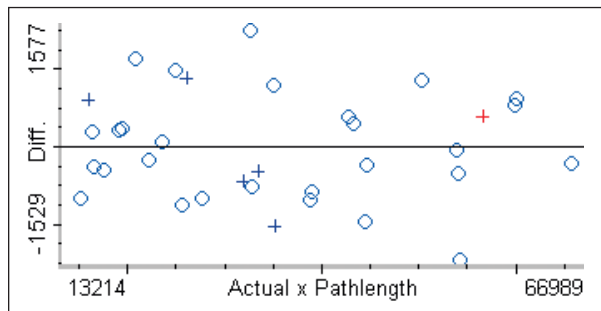


Figure 8: Plot showing prediction error versus concentration of phosphorus for the validation sample set

Since phosphorus is the more difficult analysis, a graphical presentation of the validation data fit to the calibration is presented in Figure 8. These data show a better fit than the quoted 0.17% error of prediction number in Table 2 because they are determined by sister samples to the calibration set. The numbers in Table 2 represent separate production samples.

## Conclusion

Application of PLS to the analysis of BPSG films makes it possible for a single test instrument to analyze film thickness, and boron and phosphorus concentration, in a single measurement. Computer-controlled wafer handling and mapping hardware also provides an integrated, accurate, multi-point film analysis.

Statistical quantitative analysis methods such as PLS are valuable when analyzing difficult samples like BPSG films. PLS is also useful for looking at different dielectrics and dopant concentrations. Different implementations are being studied to improve the prediction capability.

## Bibliography

1. Haaland, D.M., “Quantitative Infrared Analysis of Borophosphosilicate Films Using Multivariate Statistical Methods,” *Analytical Chemistry*, Vol. 60 No. 11 (1988).
2. Kern, Werner, “Thin Films Technology for Advanced Semiconductors: Part 2b-Deposited Dielectrics for VLSI,” *Semiconductor International* (July 1985).
3. Hurley, K.H., “Process Conditions Affecting Boron and Phosphorus in Borophosphosilicate Glass Films as Measured by FTIR Spectroscopy,” *Solid State Technology* (March 1987).
4. Tenney, A.S., and Ghezzi, M., “Composition of Phosphosilicate Glass by Infrared Absorption,” *J. Electrochem. Soc.*, Vol. 120, No. 9 (September 1973).
5. Fuller, M.P., Ritter, G.L., and Draper, C.S., “Partial Least-Squares Quantitative Analysis of Infrared Spectroscopic Data. Part 1: Algorithm Implementation,” *Applied Spectroscopy*, Volume 42, Number 2 (1988).
6. Fuller, M.P., Ritter, G.L., and Draper, C.S., “Partial Least-Squares Quantitative Analysis of Infrared Spectroscopic Data. Part 2: Application to Detergent Analysis,” *Applied Spectroscopy*, Volume 42, Number 2 (1988).
7. Tong, T.E., Schertenleib, K., and Carpio, R.A., “Process and Film Characterization of PECVD Borophosphosilicate Films for VLSI Applications,” *Solid State Technology* (January 1984).
8. Adams, A.C., “Plasma Deposition of Inorganic Films,” *Solid State Technology* (April 1983).
9. McClure, G., ed. “Computerized Quantitative Infrared Analysis,” ASTM STP 934 (1987).

## Footnotes

1. AN50637 FT-IR Measurement of PSG and BPSG Films, Part I: IR Principles and K-matrix Approach.

In addition to these offices, Thermo Fisher Scientific maintains a network of representative organizations throughout the world.

**Africa**  
+43 1 333 5034 127

**Australia**  
+61 2 8844 9500

**Austria**  
+43 1 333 50340

**Belgium**  
+32 2 482 30 30

**Canada**  
+1 800 530 8447

**China**  
+86 10 8419 3588

**Denmark**  
+45 70 23 62 60

**Europe - Other**  
+43 1 333 5034 127

**France**  
+33 1 60 92 48 00

**Germany**  
+49 6103 408 1014

**India**  
+91 22 6742 9434

**Italy**  
+39 02 950 591

**Japan**  
+81 45 453 9100

**Latin America**  
+1 608 276 5659

**Middle East**  
+43 1 333 5034 127

**Netherlands**  
+31 76 579 55 55

**South Africa**  
+27 11 570 1840

**Spain**  
+34 914 845 965

**Sweden/Norway/Finland**  
+46 8 556 468 00

**Switzerland**  
+41 61 48784 00

**UK**  
+44 1442 233555

**USA**  
+1 800 532 4752

[www.thermo.com](http://www.thermo.com)



Thermo Electron Scientific Instruments LLC, Madison, WI USA is ISO Certified.

AN50638\_E 04/08M