

# Use of Factorial Experimental Design in Static and Dynamic Light Scattering Characterization of Water-Soluble Polymers

R. D. HESTER\* and L. M. FLESHER

Department of Polymer Science, University of Southern Mississippi, Box 10076, Hattiesburg, Mississippi 39406-0076

## SYNOPSIS

When characterizing solutions of random coil polymers by static light scattering (SLS) or dynamic light scattering (DLS), linear regression is used to fit experimental data to theoretical relationships. These relationships are expressed as polynomial equations, which contain two independent variables—sample concentration and scattering angle—and a response or dependent variable that is related to radiation intensities (SLS) or intensity fluctuations (DLS). The coefficients of the terms in the polynomial are used to estimate parameters such as molecular weight and polymer coil radius of gyration. One major problem during data analysis involves deciding which polynomial model is appropriate for use with the data that inherently contains a high level of random noise that is produced by the presence of dust in the solutions. Dust is an especially troublesome problem when dealing with large polymers in aqueous solutions. Polynomial models having more terms than justified are unacceptable because the coefficients of these models are excessively corrupted by the noise. Thus, conclusions from unjustified models can be erroneous. This article discusses use of a factorial experimental design technique that obtains an acceptable model for fitting light scattering data containing high levels of random noise. © 1995 John Wiley & Sons, Inc.

## INTRODUCTION

Light scattering has been used extensively in the past to characterize high molecular weight, water-soluble polymers.<sup>1-4</sup> After dust-free dilute polymer solutions are obtained by filtration or centrifugation, collection of scattering data is very fast and can be automated. Light scattering has no upper molecular weight limit and can be used for both organic soluble as well as aqueous soluble macromolecules. Most modern instruments are capable of performing both static and dynamic light-scattering experiments.

When analyzing information on polymer solutions from both static light-scattering (SLS) and dynamic or quasielastic light-scattering (DLS) experimentation, linear regression is used to fit data to theoretical relationships. These relationships are usually expressed as polynomial equations. These equations contain two independent variables—sample concentration and scattering angle—and a response or dependent variable related to radiation

intensities, as is the case for SLS, or apparent translational diffusion coefficients for DLS. The coefficients of the polynomials can be related to macromolecular parameters such as molecular weight, coil radius of gyration, and solvent-polymer interaction. Thus, accurate calculation of the polynomial coefficients is essential for proper interpretation of the experimental data.

A major difficulty in data analysis involves determining which polynomial model is appropriate when dealing with real data that is inherently imprecise because it contains random experimental noise. The noise level can be especially high for aqueous polymer solutions that tend to attract and retain dust. Also, as the average polymer size and/or molecular weight distribution increases, solution cleanup by filtration or centrifugation becomes limited because the polymer molecules in solution and dust particles approach the same size. Thus, the experimenter is forced to deal with the high uncertainty always introduced into the scattering data by a large level of noise. Under these situations, a statistical approach must be applied to define the appropriate polynomial model that can be used.

\* To whom all correspondence should be addressed.

As the noise level increases, statistical arguments can be used to show that only lower degree polynomials can be justified unless very large quantities of data are taken or the level of confidence for estimating coefficients is lowered to unacceptable values. Usually, confidence levels are set in the 90 to 95% range to diminish both Type I and Type II errors.<sup>5</sup> A Type I error is one that rejects a model when it should be accepted, and a Type II error is one that accepts a model when it should be rejected. Models having more terms than justified by the data are corrupted because they are fitting excess noise into the model. In this case, the model will lead the experimenter to erroneous conclusions.

This article discusses the analysis of light-scattering data by using a two-variable, four-level factorial experimental design technique that employs the advantages of orthogonal polynomials. Surface analysis of the polynomial model that was justified by the experimental design will also be done using a contour plotting technique originally suggested by Zimm.<sup>6</sup>

## LIGHT-SCATTERING TEST MODEL

In the case of light scattering, a good model for the experiment is a polynomial equation of the form

$$R = B_0 + B_1X + B_2X^2 + B_3XY + B_4Y + B_5Y^2 \quad (1)$$

In eq. (1),  $R$  is the measured response and is dependent upon the independent variables  $X$  and  $Y$ , which are set by the experimenter. Note that this polynomial is second order with respect to both independent variables, and also has a term that accounts for possible interaction between the two variables. Thus, the surface described by eq. (1) can be curved. The  $B$  parameters are the coefficients that are to be estimated by fitting eq. (1) to the experimental data obtained from a set of test conditions. In light scattering,  $X$  is the square of the sin of half the scattering angle,  $\theta$ , and  $Y$  is the polymer concentration,  $C$ , in the solution scattering the radiation.

$$X = \sin^2(\theta/2) \quad (2)$$

$$Y = C \quad (3)$$

The response variable,  $R$ , measured at each test condition, depends upon the type of scattering experiment. For SLS

$$R = \frac{KC}{R_\theta} \quad (4)$$

where  $K$  is an optical constant and  $R_\theta$  is the Rayleigh ratio, which is a measure of the intensity of the scattered radiation at angle  $\theta$ . For DLS,  $R$  is the apparent translational diffusion coefficient,  $D_{app}$ , measured at the test condition. It is determined from a digital autocorrelation of the scattered radiation intensity variation that is produced by polymer center of mass motion.

$$R = D_{app} \quad (5)$$

The test model proposed for light scattering is consistent with theoretical expectations for polydispersed linear macromolecules that behave as random coils in solution. However, some terms in the test model may not be justified, depending upon the polymer-solvent system under study and the instrument's capabilities. Indeed, in some unusual cases, such as use of abnormally high solution concentrations or high scattering angles, the test model may need additional higher order terms. However, in most light scattering experiments using dilute solutions, the proposed polynomial test model should be adequate. After regression, all  $B$  coefficients not justified in the test model will be set to zero. In most cases, a simpler model, having fewer terms, will evolve after fitting the test model to the scattering data. Fitting of the test model by linear regression of data obtained from an orthogonal factorial test design is an efficient mathematical technique to estimate the values of all coefficients in the test model.

## ORTHOGONAL FACTORIAL TEST DESIGN

When using light scattering to characterize a polymer, it is convenient to measure responses using four sample concentrations. At each sample concentration, four scattering angles are used. Thus, 16 test conditions are established when characterizing a polymer solution by light scattering.

The scattering can usually be conducted such that both sample concentrations and angles of measurement are equally spaced so that a factorial experimental design can be performed.<sup>7</sup> Independent variables can then be scaled or transformed into a coded space. For example, if  $X$  and  $Y$  are varied by spacings  $2\Delta X$  and  $2\Delta Y$ , respectively, such that four levels exist for each variable, then the coded space variables,  $x$  and  $y$ , can be defined by eqs. (6) and (7). Each coded independent variable will have four values,  $-3$ ,  $-1$ ,  $1$ , and  $3$ , which represent the four test

**Table I** 4<sup>2</sup> Experimental Factorial Design

Test Condition Number	Coded Level for Independent Variable $x$	Coded Level for Independent Variable $y$
1	-3	-3
2	-3	-1
3	-3	1
4	-3	3
5	-1	-3
6	-1	-1
7	-1	1
8	-1	3
9	1	-3
10	1	-1
11	1	1
12	1	3
13	3	-3
14	3	-1
15	3	1
16	3	3

condition levels, low, middle low, middle high, and high, respectively.

$$x = \frac{X - \bar{X}}{\Delta X} \quad (6)$$

$$y = \frac{Y - \bar{Y}}{\Delta Y} \quad (7)$$

In the equations above,  $\bar{X}$  and  $\bar{Y}$  are the averages of the four  $X$  and four  $Y$  values, respectively.

We can now write the following coded space test model for each of the 16 test conditions.

$$R_i = b_0 + b_1x_i + b_2x_i^2 + b_3x_iy_i + b_4y_i + b_5y_i^2 \quad (8)$$

We can center eq. (8) by subtracting the average of all test condition responses,  $\bar{R} = b_0 + 5b_2 + 5b_5$ . This gives

$$R_i - \bar{R} = b_1x_i + b_2(x_i^2 - 5) + b_3x_iy_i + b_4y_i + b_5(y_i^2 - 5) \quad (9)$$

We can use eq. (9) to write 16 equations that describe the 16 test conditions. These equations can then be arranged into a 4<sup>2</sup> factorial design (two independent variables each having four levels) to form a set of orthogonal polynomial equations. Table I shows how the test conditions should be arranged to obtain a set of equations having orthogonal properties. Linear regression can then be performed on this set of equations to estimate the values of the

coded coefficients in the scaled and centered test model, eq. (9).

Higher order factorial designs could be used to improve data analysis when dealing with samples scattering light in a highly nonlinear manner. For example, a 5<sup>2</sup> or 6<sup>2</sup> design could be used; however, the number of test conditions or experimental effort increases to the square of the number of levels.

## LINEAR REGRESSION IN CODED SPACE

Matrix algebra can be used on the set of orthogonal equations formed by eq. (9) to find the vector  $\underline{b}$ , which has the estimated values of coded space coefficients  $b_1$  through  $b_5$ . Coefficient  $b_0$  can be found by recalling that  $b_0 = \bar{R} - 5(b_2 + b_5)$ . The matrix operation to find the vector  $\underline{b}$  is given by

$$\underline{b} = (\underline{M}^T \underline{M})^{-1} \underline{M}^T (\underline{R} - \bar{R}) \quad (10)$$

where  $\underline{R}$  is the response vector,  $\underline{M}$  is the matrix of coded test conditions,  $\underline{M}^T$  is the transpose of  $\underline{M}$  and  $(\underline{M}^T \underline{M})^{-1}$  is the inverse of the product of  $\underline{M}^T$  and  $\underline{M}$ . The matrix  $\underline{M}$  is shown in Table II. Each of the rows of the matrix  $\underline{M}$  describes one of the 16 test conditions. Each column of a row, starting with the second column to the left, gives values of  $x$ ,  $(x^2 - 5)$ ,  $xy$ ,  $y$ , and  $(y^2 - 5)$  for that test condition.

## CONFIDENCE INTERVAL FOR CODED SPACE TEST MODEL COEFFICIENTS

The variance of the six coefficients in the vector  $\underline{b}$  can be determined if the experimental standard error

**Table II** Matrix  $M$ 

	$x$	$x^2 - 5$	$xy$	$y$	$y^2 - 5$
1	-3	4	9	-3	4
1	-3	4	3	-1	-4
1	-3	4	-3	1	-4
1	-3	4	9	3	4
1	-1	-4	3	-3	4
1	-1	-4	1	-1	-4
1	-1	-4	-1	1	-4
1	-1	-4	-3	3	4
1	1	-4	-3	-3	4
1	1	-4	-1	-1	-4
1	1	-4	1	1	-4
1	1	-4	3	3	4
1	3	4	-9	-3	4
1	3	4	-3	-1	-4
1	3	4	3	1	-4
1	3	4	9	3	4

associated with the test conditions,  $s_e$ , can be estimated. Replications of a test condition or conditions can be used to estimate  $s_e$ . If  $g$  test conditions are truly replicated, then the variance of each set of  $z$  measurements at each test condition,  $v$ , having  $z - 1$  degrees of freedom, can be pooled to find an estimate of  $s_e$  by the following relationship.<sup>8</sup>

$$s_e^2 = \frac{\sum_{i=1}^g (v_i)(z_i - 1)}{\sum_{i=1}^g (z_i - 1)} \quad (11)$$

The elements on the diagonal of the variance-covariance matrix,  $\underline{V}$ , contain the variances for the coefficients of vector  $\underline{b}$ . This matrix can be evaluated by the following operation

$$\underline{V} = (\underline{M}^T \underline{M})^{-1} s_e^2 \quad (12)$$

For example, the variances of the coefficients  $b_1$  and  $b_2$ , which are designated  $v_1$  and  $v_2$  are matrix elements  $\underline{V}_{1,1}$  and  $\underline{V}_{2,2}$ , respectively. Because of the nature of an orthogonal experimental design, the covariance terms (elements not on the diagonal) in the matrix  $\underline{V}$  will be zero. No variance exists due to interactions between coefficients. Thus, the estimated standard error of each coefficient,  $s_j$ , where  $j$  varies from 1 to 5, is the square root of its variance found in matrix  $\underline{V}$ . However, because  $b_0$  is calculated from the values of  $b_2$  and  $b_5$ , its variance,  $v_0$ , must be calculated as  $v_0 = 25v_2 + 25v_5 + \underline{V}_{0,0}$ .

A Student  $t$  distribution can now be used to make confidence limits on each coefficient. The  $(1 - \alpha)$  confidence limits for each of the six coefficients are given by  $(b_j \pm t_{\alpha/2} s_j)$ . A good value to use for the significance,  $\alpha$ , is 0.10 (90% confidence). The tabulated value for  $t_{0.05}$  at 10 degrees of freedom (16 test conditions less 6 fitted coefficients gives 10 degrees of freedom) is 1.81.

If the confidence limit for a coefficient overlaps with zero, that coefficient cannot be justified and its value can be set to zero. Thus, this  $t$ -test eliminates unjustified terms in the coded test model and prevents overfitting of the data to a model having too many parameters. If  $b_2$  or  $b_5$  is eliminated (set to zero) by the  $t$ -test, then the value for  $b_0$  and its variance,  $v_0$ , should be recalculated. Recall that  $b_0 = \bar{R} - 5(b_2 + b_5)$  and  $v_0 = 25v_2 + 25v_5 + \underline{V}_{0,0}$ .

### ADJUSTED CORRELATION COEFFICIENT

After the  $t$ -test, a measure of how good the coded model fits the experimental data can be made by

calculating the correlation coefficient,  $r$ .<sup>9</sup> As the fit improves, the value of  $r$  will approach one. If  $\underline{R}$  is the response vector containing  $m$  values ( $m = 16$ ) and having an average of  $\bar{R}$ , then

$$r = \sqrt{\frac{(\underline{b}^T \underline{M}^T \underline{R} - m\bar{R}^2)}{(\underline{R}^T \underline{R} - m\bar{R}^2)}} \quad (13)$$

### COEFFICIENTS OF THE REAL-SPACE TEST MODEL

After the values of the coefficients in the coded test model have been determined, then the coefficients for the real-space test model can be calculated from the following relationships.

$$B_0 = b_0 - \frac{b_1 \bar{X}}{\Delta X} + \frac{b_2 \bar{X}^2}{\Delta X^2} + \frac{b_3 \bar{X} \bar{Y}}{\Delta X \Delta Y} - \frac{b_4 \bar{Y}}{\Delta Y} + \frac{b_5 \bar{Y}^2}{\Delta Y^2} \quad (14)$$

$$B_1 = \frac{b_1}{\Delta X} - \frac{2b_2 \bar{X}}{\Delta X^2} - \frac{b_3 \bar{Y}}{\Delta X \Delta Y} \quad (15)$$

$$B_2 = \frac{b_2}{\Delta X^2} \quad (16)$$

$$B_3 = \frac{b_3}{\Delta X \Delta Y} \quad (17)$$

$$B_4 = \frac{b_4}{\Delta Y} - \frac{2b_5 \bar{Y}}{\Delta Y^2} - \frac{b_3 \bar{X}}{\Delta X \Delta Y} \quad (18)$$

$$B_5 = \frac{b_5}{\Delta Y^2} \quad (19)$$

After substitution of the values for coefficients  $B_0$  through  $B_5$ , eq. (1) can be evaluated using a surface analysis technique originally developed by Zimm.

### STANDARD ERROR OF REAL-SPACE TEST MODEL COEFFICIENTS

The variance,  $v_u$  of a dependent variable,  $u$ , which is a known function of  $m$  independent variables,  $w_1, w_2, \dots, w_m$ , can be estimated from the variance of the independent variables  $v_{wi}$ , and the known function,  $f$ . If  $u = f(w_1, w_2, \dots, w_m)$  then the variance of  $u$  is given by<sup>15</sup>

$$v_u = \sum_{i=1}^m \left( \frac{\partial f}{\partial w_i} \right)^2 v_{w_i} \quad (20)$$

The standard error of  $u$  is the square root of its variance,  $v_u$ .

We can use the above relationship to find the variances of the coefficients of the real space model, which are calculated from eqs. (14) through (19), from the variances of the coded space model coefficients,  $v_j$ . Equations (14) through (19) are the functions from which the partial derivatives of eq. 20 can be determined. Recall that the variances of the coded space coefficients were determined from eq. (12). If some of the coded test model coefficients are set to zero by the  $t$ -test, then the variances of these coefficients should also be set to zero. If  $b_2$  and/or  $b_5$  were set to zero, then the values of  $b_0$  and its variance,  $v_0$ , should be recalculated. Thereafter, the standard error of each real space model coefficient,  $S_j$ , can be found using eq. (20).

$$S_0 = \sqrt{v_0 + \frac{\bar{X}^2 v_1}{\Delta X^2} + \frac{\bar{X}^4 v_2}{\Delta X^4} + \frac{\bar{X}^2 \bar{Y}^2 v_3}{\Delta X^2 \Delta Y^2} + \frac{\bar{Y}^2 v_4}{\Delta Y^2} + \frac{\bar{Y}^4 v_5}{\Delta Y^4}} \quad (21)$$

$$S_1 = \sqrt{\frac{v_1}{\Delta X^2} + \frac{4\bar{X}^2 v_2}{\Delta X^4} + \frac{\bar{Y}^2 v_3}{\Delta X^2 \Delta Y^2}} \quad (22)$$

$$S_2 = \sqrt{\frac{v_2}{\Delta X^4}} \quad (23)$$

$$S_3 = \sqrt{\frac{v_3}{\Delta X^2 \Delta Y^2}} \quad (24)$$

$$S_4 = \sqrt{\frac{v_4}{\Delta Y^2} + \frac{4\bar{Y}^2 v_5}{\Delta Y^4} + \frac{\bar{X}^2 v_3}{\Delta X^2 \Delta Y^2}} \quad (25)$$

$$S_5 = \sqrt{\frac{v_5}{\Delta Y^4}} \quad (26)$$

## SURFACE ANALYSIS OF THE MODEL EQUATION

Using Zimm's technique, experimental measurements of scattering response,  $KC/R_\theta$  for SLS or  $D_{app}$  for DLS, are plotted as the ordinate vs. a compound abscissa,  $\sin^2(\theta/2) + kC$ . Use of the compound abscissa forces the points on the plot to be displaced from each other. This effect is due to the spacing constant  $k$ . The value of  $k$  is arbitrarily selected to

provide adequate distance between the experimental data points. The set of points displays the complete response grid of the scattering experiment. When a  $4^2$  factorial design is used, 16 data points will form the surface response grid.

In the original Zimm technique, this grid was used to form lines of constant angles and lines of constant concentration. The lines of constant concentration were linearly extrapolated to give a set of points that had zero concentration. Also, the lines of constant angles were linearly extrapolated to give a set of points that have zero angle. The straight line fitted to the zero concentration points and the straight line fitted to the zero angle points should intercept the ordinate at the same position. This common intercept is the reciprocal of the molecular weight for a SLS experiment or is the true translational diffusional coefficient for a DLS experiment. The slopes of the two straight lines can be used to find other polymer parameters.

Unfortunately, linear extrapolation does not work when the data points on the Zimm plot form curves instead of straight lines. This is the usual situation when analyzing high molecular weight water-soluble polymers. In this situation, the model equation developed from the  $4^2$  factorial design and justified by the  $t$ -test can be used to plot the four constant concentration curves, the four constant angle curves, the zero concentration curve, and the zero angle curve associated with the experimental data in the Zimm plot. Both curves will intercept the ordinate at the same position, and this intersection point will be equal to the model equation coefficient  $B_0$ .

The 16 intersections of the constant concentration curves with the constant angle curves are the model equation fit points that correspond to the 16 data points on the plot. Thus, a visual understanding of the model equation fit to the experimental data can be realized by noting the placement of data points relative to curve intersections. If an experimental data point exactly corresponds to an intersection of the two curves made at the angle and concentration of that data point, then a perfect fit between the model and data exists at that test condition. The greater the deviation of a data point from its corresponding intersection point, the more poorly the model fits that test condition.

Use of the above factorial design experimentation and plotting procedures can best be demonstrated by giving three examples. In Example 1, a static light experiment was performed on a water-soluble copolymer. In Example 2, the same copolymer was characterized by dynamic light scattering. Example 3 is Example 1 redone without justifying the model.

## EXPERIMENTAL

The high molecular weight random copolymer employed in the scattering experiments was supplied by C. L. McCormick, and was synthesized from acrylamide (AM) and 3-acrylamido-3-methylbutanoic acid (AMBA) monomers in the ratio of 95 to 5, as described in previous publications.<sup>10,11</sup> The aqueous solvent used to make the polymer solutions contained 0.514M NaCl and had a refractive index,  $n$ , of 1.338 at 25°C, with radiation having a wavelength,  $\lambda_0$ , of 6328 Å. The solvent was filtered using a 0.02 micron Anotop 25 inorganic membrane manufactured by Whatman. The solvent viscosity,  $\eta_0$ , was 0.934 cp. Polymer solutions were made by gentle mixing for a minimum of 4 days. Solution preparation over this time scale minimized polymer aggregation. Immediately before SLS analysis the solutions were clarified by a syringe filter using a 1.0  $\mu\text{m}$  Puradisc 25AS polysulfone membrane manufactured by Whatman.

Four solutions were made that had polymer concentrations of 0.04, 0.08, 0.12, and 0.16 g/liter. Light-scattering experiments were performed on these solutions at 25°C and at scattering angles of 32.3, 65, 90, and 115 degrees using a BI-2030AT goniometer equipped with a BI-DS photomultiplier and digital correlator manufactured by Brookhaven Instruments Corp. of Holtsville, NY. Duplicate measurements of intensity response, as  $KC/R_\theta$ , were made for each of the 16 test conditions in SLS. Triplicate measurements of the apparent translational diffusion coefficient,  $D_{\text{app}}$ , were made for each of the 16 test conditions in DLS. The  $D_{\text{app}}$  values were calculated from the autocorrelation function using the cummulant technique.<sup>19</sup>

The optical constant,  $K$ , used in SLS was calculated using a  $dn/dC$  value of 0.1559 mL/g for the polymer-solvent system under study. The change in the refractive index with respect to solution concentration,  $dn/dC$ , was measured using a KMX-16 differential refractometer manufactured by Chromatix Inc. of Sunnyvale, CA. The polymer samples were not dialyzed to equilibrium with the solvent for  $dn/dC$  measurements because this procedure is not necessary when the solvent ionic strength is high.<sup>16-18</sup>

All calculations were performed on a 80486 personal computer using the program Mathcad 5.0+ manufactured by Mathsoft Inc. of Cambridge, MA. The Zimm plots were also generated with this software using a spacing constant,  $k$ , of 5000 mL/g.

All solution and 0.514M NaCl aqueous solvent viscosity measurements were made at 25°C using a Contraves L30 rheometer. For each fluid, viscosities

were determined at 3.2, 8.1, 14.9, and 27.5 s<sup>-1</sup> shear rates and then extrapolated to find the "zero" shear fluid viscosity. The "zero" shear viscosities of four copolymer solutions containing 0.125, 0.100, 0.075, and 0.050 g/liter were then used to establish the copolymer intrinsic viscosity.

## EXAMPLE 1 (STATIC LIGHT SCATTERING)

The theoretical relationship between the properties of a polymer sample in solution and its light-scattering characteristics is expressed by the Debye relationship<sup>12</sup> given by eq. (27).

$$\frac{KC}{R_\theta} = \frac{1}{M_w} + \frac{16\pi^2 n^2 R_g^2 \sin^2\left(\frac{\theta}{2}\right)}{3M_w \lambda_0^2} - \frac{64\pi^4 n^4 R_g^4 \sin^4\left(\frac{\theta}{2}\right)}{3M_w \lambda_0^4} + 2A_2 C + 3A_3 C^2 + \text{other terms} \quad (27)$$

In eq. (27),  $M_w$  and  $R_g$  are the polymer weight average molecular weight and the "z" average polymer coil radius of gyration, respectively. The second and third virial coefficients,  $A_2$  and  $A_3$  are related to solvent-polymer interactions. Note that eq. (27) can be expressed as test model eq. (1) when we let the variables

$$R = \frac{KC}{R_\theta} \quad (28)$$

$$X = \sin^2\left(\frac{\theta}{2}\right) \quad (29)$$

$$Y = C \quad (30)$$

and let the  $B$  coefficients be defined as

$$B_0 = \frac{1}{M_w} \quad (31)$$

$$B_1 = \frac{16\pi^2 n^2 R_g^2}{3M_w \lambda_0^2} \quad (32)$$

$$B_2 = -\frac{64\pi^4 n^4 R_g^4}{3M_w \lambda_0^4} \quad (33)$$

$$B_3 = \text{interaction coefficient} \quad (34)$$

$$B_4 = 2A_2 \quad (35)$$

$$B_5 = 3A_3 \quad (36)$$

Table III Example 1 (SLS) Responses

Test Condition	Measurement 1 mol/g $\times 10^7$	Measurement 2 mol/g $\times 10^7$	Average Response mol/g $\times 10^7$
1	2.60	2.15	2.38
2	2.53	2.67	2.60
3	2.81	2.97	2.89
4	3.10	2.81	2.96
5	5.55	4.75	5.15
6	5.43	5.50	5.47
7	5.84	6.26	6.05
8	6.29	5.62	5.96
9	9.68	7.21	8.45
10	8.60	8.31	8.46
11	9.06	9.12	9.09
12	9.25	7.99	8.62
13	11.28	9.07	10.18
14	10.66	10.33	10.50
15	10.84	11.37	11.11
16	11.62	10.10	10.86

Note that we have defined the "other terms" in eq. (27) as a single expression equal to the product of an interaction coefficient,  $B_3$ , and variables  $X$  and  $Y$ .

Values for  $X$  are 0.077, 0.289, 0.500, and 0.711. Thus,  $2\Delta X = 0.211$  and  $\bar{X} = 0.394$ . Recall values for  $Y$  are 0.04, 0.08, 0.12, and 0.16 g/liter. Thus,  $2\Delta Y = 0.04$  g/liter and  $\bar{Y} = 0.10$  g/liter. Because of the equal separation of the values used for independent variables  $X$  and  $Y$ , we can use a  $4^2$  factorial design and then do a regression analysis to find estimates of the model coefficients. Thereafter we can use a  $t$ -test to justify each coefficient of the model.

Scattering experiments were performed according to the design shown by Table I. The responses,  $KC/R_\theta$  in mol/g, to the test conditions are shown in Table III. Two measurements were taken at each condition. The average at each test condition of the two

measurements was used to form the response vector,  $\bar{R}$ , and is shown in the last column of Table III. Because two response measurements were taken at each test condition, eq. (11) could be used to estimate the average experimental error,  $s_e$ , associated with a response to a test condition. The  $s_e$  value obtained was  $5.15 \times 10^{-8}$  mol per gram. The average of all responses,  $\bar{R}$ , was  $6.92 \times 10^{-7}$  mol per gram. Thus, the standard experimental error is about 7% of the average response.

The coded and scaled test model, eq. (9), can now be solved for the  $b$  coefficients using eq. (10). The results, vector  $\underline{b}$ , are shown in Table IV along with the upper and lower 90% confidence limits. The limits were calculated by solving for the matrix  $\underline{V}$  and then calculating the limits ( $b_j \pm t_{0.05} s_j$ ) after finding the  $s_j$  values from the square root of the  $\underline{V}$  diagonal elements.

Table IV Example 1 (SLS) Coded Test Model Coefficients

Coefficient	Estimated Value <sup>a</sup>	Standard Error <sup>b</sup>	Confidence Limits	
			Upper	Lower
$b_0$	$7.34 \times 10^{-7}$	$2.61 \times 10^{-8}$	$7.81 \times 10^{-7}$	$6.87 \times 10^{-7}$
$b_1$	$1.34 \times 10^{-7}$	$5.76 \times 10^{-9}$	$1.45 \times 10^{-7}$	$1.24 \times 10^{-7}$
$b_2$	$-5.90 \times 10^{-9}$	$3.22 \times 10^{-9}$	$-7.45 \times 10^{-11}$	$-1.17 \times 10^{-8}$
$b_3$	$1.63 \times 10^{-11}$	$2.57 \times 10^{-9}$	$4.68 \times 10^{-9}$	$-4.64 \times 10^{-9}$
$b_4$	$1.11 \times 10^{-8}$	$5.76 \times 10^{-9}$	$2.15 \times 10^{-8}$	$6.51 \times 10^{-10}$
$b_5$	$-2.52 \times 10^{-9}$	$3.22 \times 10^{-9}$	$3.30 \times 10^{-9}$	$-8.35 \times 10^{-9}$

$$^a b_0 = \bar{R} - 5(b_2 + b_5).$$

$$^b s_0 = \sqrt{25v_2 + 25v_5 + V_{\alpha}}.$$

Inspection of Table IV shows that coefficients  $b_3$  and  $b_5$  are not significantly different from zero and, thus, their values and variances will be set to zero in subsequent calculations. Because  $b_5$  was set to zero,  $b_0$  and  $s_0$  were recalculated as  $7.21 \times 10^{-7}$  and  $2.06 \times 10^{-8}$ , respectively. Equations (14) through (19) can now be used to find the test model  $B$  coefficients. Next eqs. (31), (32), and (34) can be used to find the polymer parameters  $M_w$ ,  $R_g$ , and  $A_2$ .

Equation (20) can be used to find the standard error of the polymer parameters from the functions described by eqs. (31), (32), and (35). If  $S_{M_w}$ ,  $S_{R_g}$ , and  $S_{A_2}$  are the standard error of the molecular weight, radius of gyration and second virial coefficient, respectively, then

$$S_{M_w} = \frac{S_0}{B_0^2} \quad (37)$$

$$S_{A_2} = \frac{S_4}{2} \quad (38)$$

$$S_{R_g} = \frac{\lambda_0}{8\pi n} \sqrt{3B_0 B_1 S_{M_w}^2 + \frac{3S_1^2}{B_0 B_1}} \quad (39)$$

The  $B$  coefficients, the polymer parameters, and all standard errors are listed in Table V.

Figure 1, a Zimm plot, shows the average response for each test condition as "X" symbols, which are superimposed onto the curves constructed from the test model. Curves of constant angle are shown solid, curves of constant concentration are shown dashed, and the extrapolated zero concentration and zero angle curves are both shown dotted. The plot shows that the fit of the model equation to the experimental data is adequate. The adjusted correlation coefficient,  $r$ , calculated from eq. (13), has a value of 0.997, which also indicates that the model gives a good fit to the data.

## EXAMPLE 2 (DYNAMIC LIGHT SCATTERING)

The apparent translational diffusional coefficient,  $D_{app}$ , obtained by DLS is related to the true diffusional coefficient,  $D_{true}$ , by the relationship

$$D_{app} = D_{true} + \alpha \sin^2\left(\frac{\theta}{2}\right) + \beta \sin^4\left(\frac{\theta}{2}\right) + \chi C \sin^2(\theta/2) + \delta C + \xi C^2 \quad (40)$$

The coefficients,  $D_{true}$ ,  $\alpha$ ,  $\beta$ ,  $\chi$ ,  $\delta$ ,  $\xi$ , are related to polymer-solvent properties. The parameters  $\delta$  and

$\xi$  are usually referred to as the second and third diffusional virial coefficients. They are related to the interactions between solvent and polymer. The parameter  $D_{true}$  is the diffusional coefficient in the limit of zero polymer concentration and zero scattering angle, and is related to the hydrodynamic polymer coil radius,  $R_h$ , by the Stokes-Einstein equation.

$$R_h = \frac{k_B T}{6\pi\eta_0 D_{true}} \quad (41)$$

In eq. (41),  $k_B$  is the Boltzmann constant,  $T$  is the absolute temperature, and  $\eta_0$  is the solvent viscosity. For random coil polymers in solution, the coefficients  $\alpha$  and  $\beta$  are expected to be related to  $D_{true}$  and  $R_g$  by the following relationships:<sup>13</sup>

$$\alpha = \frac{16\pi^2 n^2 D_{true} R_g^2}{5\lambda_0^2} \quad (42)$$

$$\beta = \frac{-128\pi^4 n^4 D_{true} R_g^4}{5\lambda_0^4} \quad (43)$$

The coefficients  $\delta$  and  $\xi$  are related to  $D_{true}$  and the polymer intrinsic viscosity,  $[\eta]$ .

$$\delta = \kappa D_{true} [\eta] \quad (44)$$

$$\xi = \nu D_{true} [\eta]^2 \quad (45)$$

The proportionality constants,  $\kappa$  and  $\nu$ , are expected to have values of 1.56 and -6.83, respectively.<sup>14</sup> Jamieson<sup>1</sup> has suggested that  $\kappa$  may have a lower value of 0.58.

Equation (40) can be expressed as test model eq. (1) when we let the variables be defined as

$$R = D_{app} \quad (46)$$

$$X = \sin^2\left(\frac{\theta}{2}\right) \quad (47)$$

$$Y = C \quad (48)$$

and let the  $B$  coefficients be defined as

$$B_0 = D_{true} \quad (49)$$

$$B_1 = \alpha \quad (50)$$

$$B_2 = \beta \quad (51)$$

$$B_3 = \chi \quad (52)$$

$$B_4 = \delta \quad (53)$$

$$B_5 = \xi \quad (54)$$



Table V Example 1 (SLS) Test Model Coefficients and Polymer Parameters

Coefficient	Estimated Value	Standard Error	Polymer Parameter	Estimated Value	Standard Error
$B_0$	$8.3 \times 10^{-8}$	$6.3 \times 10^{-8}$	Molecular weight, $M_w$	$1.2 \times 10^7$ g/mol	$0.9 \times 10^7$ g/mol
$B_1$	$1.7 \times 10^{-6}$	$2.4 \times 10^{-7}$	Radius of gyration, $R_g$	2900 Å	1100 Å
$B_2$	$-5.3 \times 10^{-7}$	$2.9 \times 10^{-7}$	—	—	—
$B_3$	0	—	—	—	—
$B_4$	$5.5 \times 10^{-4}$	$2.9 \times 10^{-4}$	Second virial coefficient, $A_2$	$2.7 \times 10^{-4}$ mL mol/g <sup>2</sup>	$1.4 \times 10^{-4}$ mL mol/g <sup>2</sup>
$B_5$	0	—	—	—	—

Values for  $X$ ,  $2\Delta X$ ,  $\bar{X}$ ,  $Y$ ,  $2\Delta Y$ , and  $\bar{Y}$  are the same as in the example for SLS and, thus, a  $4^2$  factorial design and  $t$ -test can be applied as in Example 1. The DLS responses,  $D_{app}$  values in  $\text{cm}^2/\text{s}$ , are shown in Table VI. Three measurements were taken at each of the 16 test conditions. The average at each test condition of the three measurements was used to form the response vector,  $\bar{R}$ , and is shown in the last column of Table VI. As in the SLS example, an experimental error,  $s_e$ , was calculated to be  $1.47 \times 10^{-9} \text{ cm}^2/\text{s}$ . The average response,  $\bar{R}$ , was  $3.13 \times 10^{-8} \text{ cm}^2/\text{s}$ . Thus, the standard experimental error is about 5% of the average response.

The coded and scaled test model, eq. (9), can now be solved for the  $b$  coefficients using eq. (10). The results, vector  $\underline{b}$ , are shown in Table VII along with the upper and lower 90% confidence limits calculated by solving for the matrix  $\underline{V}$  and then calculating the limits  $b_j \pm t_{0.05} s_j$  after finding the  $s_j$  values from the  $\underline{V}$  diagonal elements.

Inspection of Table VII shows that coefficients  $b_3$ ,  $b_4$ , and  $b_5$  are not significantly different from zero and, thus, their values and variances will be set to zero in subsequent calculations. Because  $b_5$  was set to zero,  $b_0$  and  $s_0$  were recalculated as  $3.39 \times 10^{-8}$  and  $5.86 \times 10^{-10}$ , respectively. Equations (14)

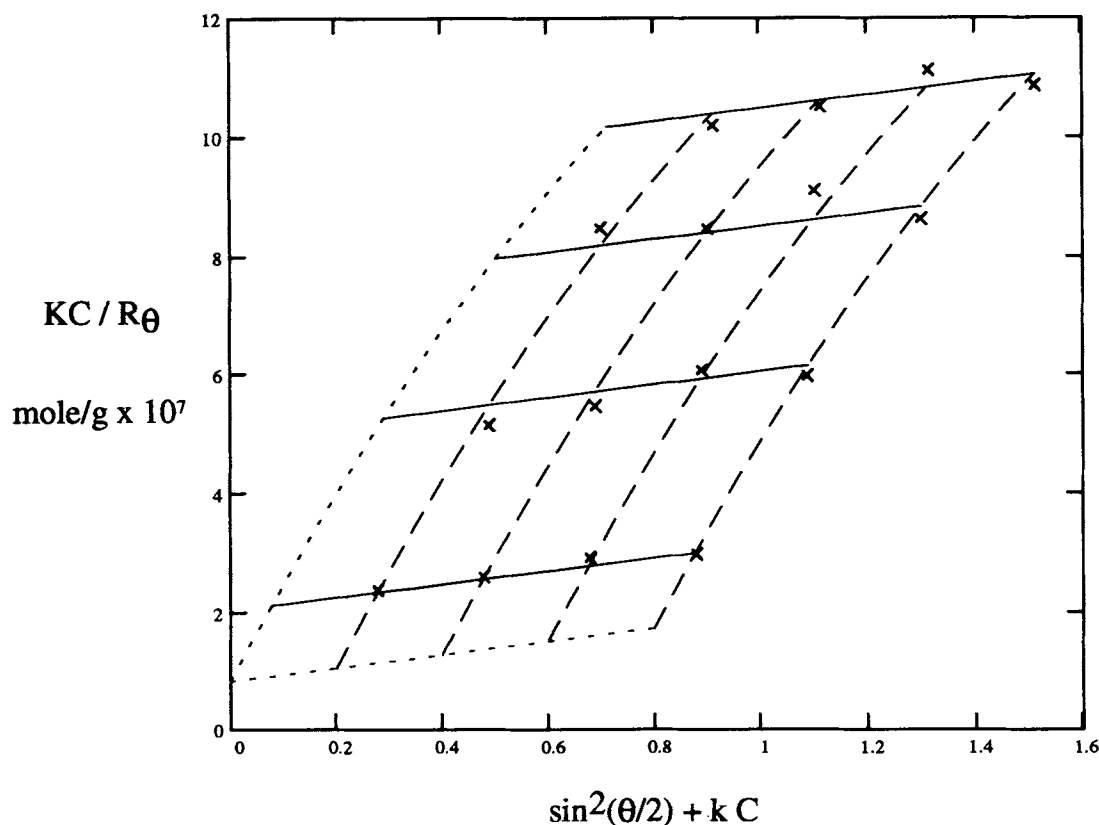


Figure 1 Zimm plot for Example 1 (SLS).

**Table VI Example 2 (DLS) Responses**

Test Condition	Measurement 1 cm <sup>2</sup> /s × 10 <sup>8</sup>	Measurement 2 cm <sup>2</sup> /s × 10 <sup>8</sup>	Measurement 3 cm <sup>2</sup> /s × 10 <sup>8</sup>	Average Response cm <sup>2</sup> /s × 10 <sup>8</sup>
1	1.95	2.05	1.85	1.95
2	1.81	1.75	1.81	1.79
3	1.36	1.79	1.84	1.67
4	1.88	1.91	1.83	1.87
5	3.04	3.02	2.96	3.00
6	3.01	3.00	2.97	2.98
7	3.09	3.02	2.97	3.03
8	2.85	3.00	2.88	2.91
9	3.80	3.80	3.77	3.79
10	3.74	3.65	3.72	3.71
11	3.62	3.68	3.60	3.63
12	3.65	3.64	3.62	3.63
13	3.92	4.10	4.18	4.07
14	3.61	4.18	4.28	4.03
15	4.14	4.24	4.17	4.18
16	4.11	3.23	4.13	3.82

through (19) can now be used to find the test model  $B$  coefficients. Equations (49) and (41) can be used to find  $D_{\text{true}}$  and the hydrodynamic polymer coil radius,  $R_h$ . The polymer coil radius of gyration,  $R_g$ , can be calculated from eq. 42.

As in Example 1, eq. (20) can be used to find the standard error of the polymer parameters from the functions described by eqs. (41), (42), and (49). If  $S_D$ ,  $S_{RG}$ , and  $S_{Rh}$  are the standard error of the true diffusion coefficient, dynamic radius of gyration and hydrodynamic radius, respectively, then

$$S_D = S_0 \quad (54)$$

$$S_{Rh} = \left( \frac{k_B T}{6\pi\eta_0} \right) \frac{S_0}{B_0^2} \quad (55)$$

$$S_{RG} = \frac{\lambda_0}{8\pi n} \sqrt{\frac{5S_1^2}{B_0 B_1} + \frac{5B_1 S_0^2}{B_0^3}} \quad (56)$$

The  $B$  coefficients, the polymer parameters, and all standard errors are listed in Table VIII. Figure 2 is the Zimm plot for the DLS example. The fit of experimental data is shown to be adequate. The adjusted correlation coefficient,  $r$ , calculated from eq. (13), has a value of 0.994, which also indicates that the model gives a good fit to the data.

### EXAMPLE 3 (SLS WITH UNJUSTIFIED MODEL)

If data analysis is done without performing a  $t$ -test to remove unjustified terms in the model, results

can be misleading. This is demonstrated by Example 3, which is a reanalysis of the SLS data of Example 2 without using model justification. In this example, all terms of the coded model are retained after regression. The Zimm plot, Figure 3, shows an excellent fit of the unjustified model to the data. Also, the adjusted correlation coefficient has a value of 0.998, which indicates a better fit than the model of Example 1. However, the molecular parameters calculated for Example 3 ( $M_w = 3.1 \times 10^7$  g/mol,  $R_g = 4700$  Å, and  $A_2 = 9.1 \times 10^{-4}$  mL mol/g<sup>2</sup>) are significantly larger than determined in Example 1 ( $M_w = 1.2 \times 10^7$  g/mol,  $R_g = 2900$  Å, and  $A_2 = 2.7 \times 10^{-4}$  mL mol/g<sup>2</sup>). These molecular parameters are too large to be consistent with the intrinsic viscosity measured for this polymer-solvent system, 17.0 dL/g. Thus, without model justification, erroneous conclusions can be obtained when analyzing light scattering data.

### CONCLUSIONS

The data analysis technique detailed above, which employs a factorial experimental design, can be used to establish a light-scattering model that is statistically justified. Calculations are straightforward and can be easily programmed. The model can thereafter be used to estimate macromolecular parameters such as polymer coil size, viral and diffusional coefficients, and molecular weight. In addition, confidence intervals can be found for the polymer parameters estimated from the model.

Table VII Example 2 (DLS) Coded Test Model

Coefficient	Value <sup>a</sup>	Standard Error <sup>b</sup>	Confidence Limits	
			Upper	Lower
$b_0$	$3.38 \times 10^{-8}$	$7.44 \times 10^{-10}$	$3.52 \times 10^{-8}$	$3.25 \times 10^{-8}$
$b_1$	$3.66 \times 10^{-9}$	$1.64 \times 10^{-10}$	$3.96 \times 10^{-9}$	$3.37 \times 10^{-9}$
$b_2$	$-5.18 \times 10^{-10}$	$9.16 \times 10^{-11}$	$-3.52 \times 10^{-10}$	$-6.83 \times 10^{-10}$
$b_3$	$-2.49 \times 10^{-11}$	$7.33 \times 10^{-11}$	$1.08 \times 10^{-10}$	$-1.58 \times 10^{-10}$
$b_4$	$-2.15 \times 10^{-10}$	$1.64 \times 10^{-10}$	$8.18 \times 10^{-11}$	$-5.11 \times 10^{-10}$
$b_5$	$5.57 \times 10^{-12}$	$9.16 \times 10^{-11}$	$1.71 \times 10^{-10}$	$-1.60 \times 10^{-10}$

$$^a b_0 = \bar{R} - 5(b_2 + b_5).$$

$$^b s_0 = \sqrt{25v_2 + 25v_5 + V_{\bar{R}}}.$$

We wish to thank the Department of Energy (contract number DE-AC22-92BC 14882) for its support of this work and Hartono Halim for his assistance in obtaining the light-scattering experimental data.

## NOMENCLATURE

$A_2$	Second virial coefficient
$A_3$	Third virial coefficient
$B_j$	Coefficients of the test model, eq. (1)
$b_j$	Coefficients of the coded test model, eq. (8)
$\underline{b}$	Vector containing $b_j$ coefficients
$C$	Solution concentration
$D_{\text{app}}$	Apparent translational diffusion coefficient
$D_{\text{true}}$	True translational diffusion coefficient
$g$	Index counter for eq. (11)
$i$	Index counter for experimental test conditions
$j$	Index counter for model coefficients
$K$	Light-scattering optical constant
$k$	Spacing constant for Zimm Plot
$k_B$	Boltzman constant
$\underline{M}$	Matrix defined by Table II
$\underline{M}^T$	Transpose of matrix $\underline{M}$
$\overline{M}_w$	Weight average molecular weight
$m$	Total number of test conditions
$n$	Refractive index
$R_i$	Average response at test condition $i$
$\bar{R}$	Average response of all test conditions
$\underline{R}$	Vector containing $R_i$ responses
$R_g$	Polymer radius of gyration
$R_h$	Polymer hydrodynamic radius
$R_\theta$	Rayleigh ratio of light scattering intensities
$r$	Adjusted correlation coefficient

$S_{A_2}$	Standard error of the second virial coefficient
$S_D$	Standard error of the true diffusional coefficient
$S_j$	Estimated standard error associated with real-space model coefficients
$S_{MW}$	Standard error of the weight average molecular weight
$S_{RG}$	Standard error of the radius of gyration from DLS
$S_{Rg}$	Standard error of the radius of gyration from SLS
$S_{Rh}$	Standard error of the hydrodynamic radius
$s_e$	Estimated experimental error
$s_j$	Estimated standard error associated with coded test model coefficients
$T$	Absolute temperature
$t_{0.05}$	Student $t$ distribution value at 90% confidence level
$u$	Dependent variable of eq. (20)
$\underline{V}$	Variance-covariance matrix
$v_j$	Variance associated with coded test model coefficients
$w_j$	Independent variables of eq. (20)
$X$	Real-space independent variable
$\bar{X}$	Average value of $X$ independent variables
$\Delta X$	Half of separation value between $X$ independent variables
$x$	Coded space independent variable
$Y$	Real-space independent variable
$\bar{Y}$	Average value of $Y$ independent variables
$\Delta Y$	Half of separation value between the $Y$ independent variables
$y$	Coded space independent variable
$z$	Number of measurements at a given test condition
$\alpha$	Coefficient in eqs. (40) and (42)

Table VIII Example 2 (DLS) Test Model Coefficients and Polymer Parameters

Coefficient	Estimated Value	Standard Error	Polymer Parameter	Estimated Value	Standard Error
$B_0$	$1.3 \times 10^{-8}$	$1.5 \times 10^{-9}$	True diffusional coefficient, $D_{\text{true}}$	$1.3 \times 10^{-8} \text{ cm}^2/\text{s}$	$1.5 \times 10^{-9} \text{ cm}^2/\text{s}$
$B_1$	$7.1 \times 10^{-8}$	$6.7 \times 10^{-9}$	Radius of gyration, $R_g$	2000 Å	150 Å
$B_2$	$-4.6 \times 10^{-8}$	$8.2 \times 10^{-9}$	Hydrodynamic radius, $R_h$	1800 Å	210 Å
$B_3$	0	—	—	—	—
$B_4$	0	—	—	—	—
$B_5$	0	—	—	—	—

$1 - \alpha$	Confidence level
$\beta$	Coefficient in eqs. (40) and (43)
$\chi$	Coefficient in eq. (40)
$\delta$	Coefficient in eqs. (40) and (44)
$\xi$	Coefficient in eqs. (40) and (45)
$\eta_0$	Solvent viscosity
$\kappa$	Proportionality constant in eq. (44)
$\lambda_0$	Wavelength of light scattering radiation
$\nu$	Proportionality constant in eq. (45)
$\theta$	Scattering angle
$dn/dC$	Change in refractive index with respect to polymer concentration

## REFERENCES

1. P. M. Patterson and A. M. Jamieson, *Macromolecules*, **18**, 266 (1985).
2. M. Skouri, J. P. Munch, S. J. Candau, S. Neyret, and F. Candau, *Macromolecules*, **27**, 69 (1994).
3. Z. Zhou and B. Chu, *Macromolecules*, **24**, 2025 (1994).
4. N. Borochov and H. Eisenberg, *Macromolecules*, **27**, 1440 (1994).
5. R. L. Wine, *Statistics for Scientists and Engineers*, Prentice-Hall, Hillsdale, NJ, 1964, p. 215.
6. B. H. Zimm, *J. Chem. Phys.*, **16**, 1099 (1948).

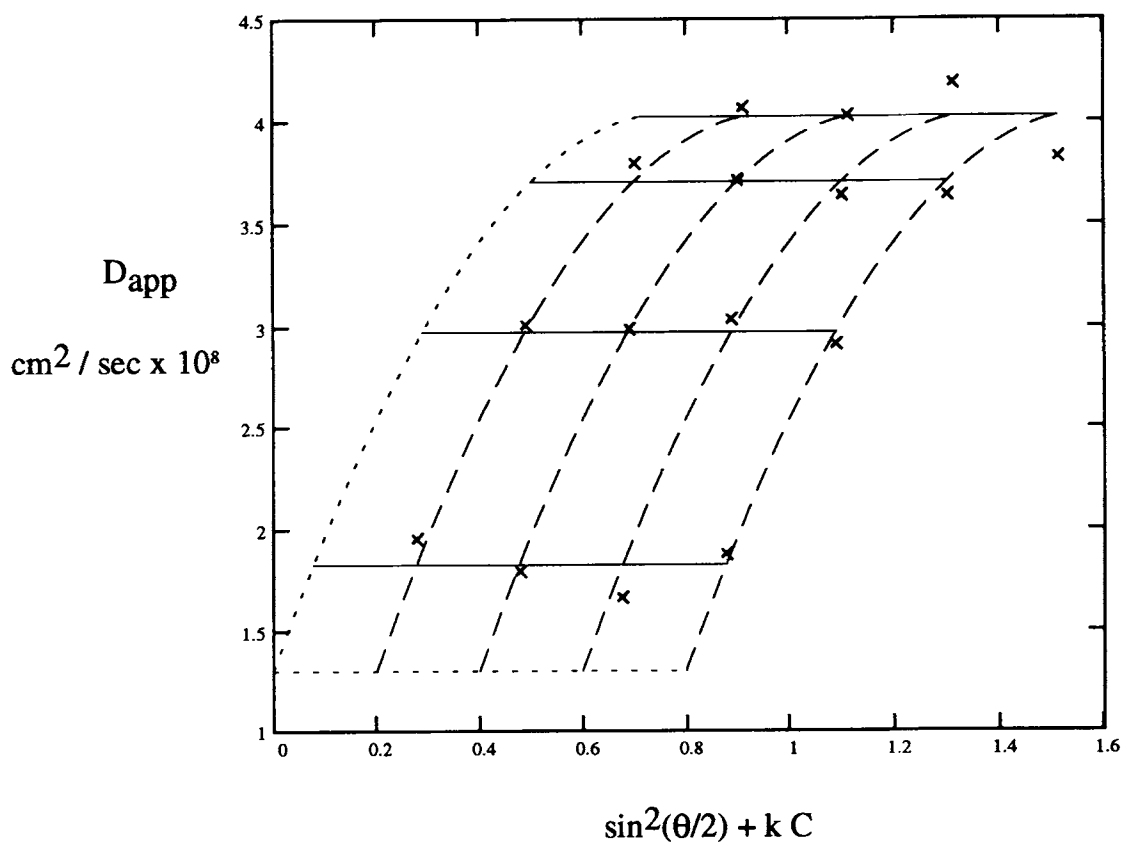


Figure 2 Zimm plot for Example 2 (DLS).

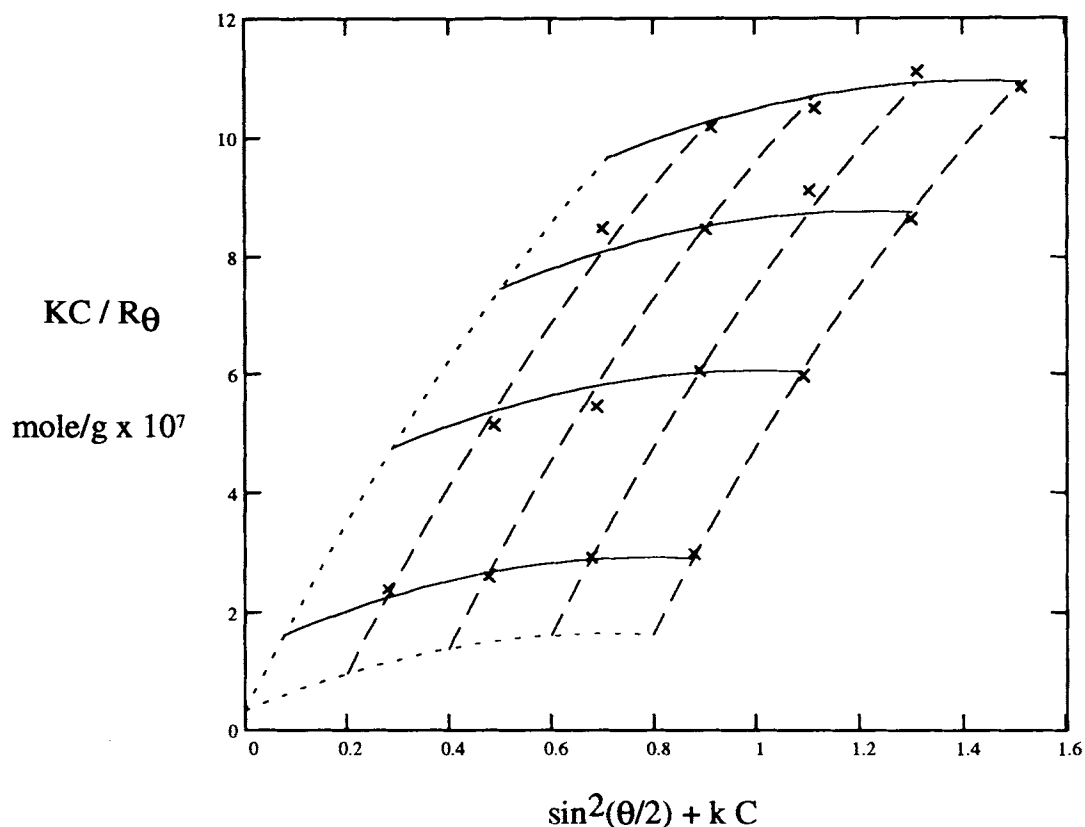


Figure 3 Zimm plot for Example 3 (SLS).

7. C. R. Hicks, *Fundamental Concepts in the Design of Experiments*, Holt, Rinehart and Winston, New York, 1964, p. 95.
8. G. P. Box, W. G. Hunter, and J. S. Hunter, *Statistics for Experiments*, John Wiley & Sons, New York, 1978, p. 319.
9. N. R. Draper and H. Smith, *Applied Regression Analysis*, 2nd Ed., John Wiley & Sons, New York, 1981, p. 91.
10. C. L. McCormick and K. P. Blackmon, *J. Polym. Sci., Polym. Chem.*, **A24**, 2635 (1986).
11. C. L. McCormick, K. P. Blackmon, and D. L. Elliott, *J. Polym. Sci., Polym. Chem.*, **A24**, 2619 (1986).
12. C. Tanford, *Physical Chemistry of Macromolecules*, John Wiley & Sons, New York, 1961, p. 275.
13. R. M. Peitzsch, M. J. Burt, and W. F. Reed, *Macromolecules*, **25**, 806 (1992).
14. K. S. Schmitz, *An Introduction to Light Scattering by Macromolecules*, Academic Press, New York, 1990, p. 181.
15. J. W. Mellor, *Higher Mathematics for Students of Chemistry & Physics*, Longman, Green & Co., London, 1926, p. 530.
16. P. Kratochvil, *Pure Appl. Chem.*, **54**(2), 379 (1982).
17. K. S. Schmitz, *An Introduction to Dynamic Light Scattering by Macromolecules*, Academic Press, New York, 1990, p. 245.
18. E. Selegny and G. Muller, *J. Polym. Sci., Part C*, **30**, 125 (1970).
19. D. E. Koppel, *J. Chem. Phys.*, **57**, 4814 (1972).

Received October 21, 1994

Accepted March 17, 1995