Factor analysis in the NMR characterization of multicomponent polymers

H. N. Cheng*, Paul C. Gillette

Hercules Incorporated Research Center, 500 Hercules Road, Wilmington, DE 19808-1599, USA

Summary

Factor analysis has been found to be suitable for the analysis NMR data of polymers which contain multiple components. Two series of ethylene/propylene copolymers have been studied by this technique. In general, the number of components that can be resolved from NMR data depends on both the quality and the quantity of available experimental data.

Introduction

Many commercial polymers are compositionally heterogeneous and contain two or more components. A multicomponent polymer may result, for example, from multi-stage reactions where different amounts of monomers are fed into a reactor at different times. Alternatively, the multiple components may be due to deliberate blending of polymers with different compositions to optimize end-use properties or to consume off-spec materials. Multicomponent polymers may also arise as a result of different kinds of initiators, catalytic sites, or propagating species that are present in the polymerizing medium. Furthermore, these may originate from phase separation that sometimes occurs during polymerization. Conformational differences arising from bond torsion variations within the amorphous phase and crystallinity represent other forms of heterogeneity. Whatever the cause(s) of this multistate heterogeneity, the resulting polymer contains components with different composition (or tacticity) and is, in general, more difficult to analyze in the context of structure/property correlations.

NMR is usually the preferred technique to study polymer microstructure. A number of NMR approaches have been devised to analyze multicomponent polymers (1-8). The analysis can be carried out on a single polymer sample, or (more preferably) on a series of related polymers. It has been pointed out (7) that four types of situations are particularly suited for a systematic NMR/computational analysis:
1. A given polymer is fractionated and the NMR data of the fractions are analyzed simultaneously.
2. A given polymerization is sampled at various conversions and the polymers analyzed simultaneously.

---

* Corresponding author

3. A series of polymers can be prepared using the same experimental procedures but with different comonomer feed ratios.
4. A series of polymers may be made under essentially the same experimental conditions whereby only a single variable is systematically varied. Examples are polymers made with a Ziegler-Natta catalyst where one experimental variable is changed that affects the activities of the different catalytic sites but does not significantly affect the nature of the sites.

A major difficulty in analyzing multicomponent polymers by NMR is to ascertain the number of components present. Sometimes, complementary or auxiliary data are available. For example, if other analytical data (e.g., SEC, TREF, and fractionation) have been collected, these may be incorporated in the analysis (8,9). Polymerization procedures, laboratory notebooks, and personal observations may also provide useful clues. However, such data are not always available or helpful. In these cases, the NMR data need to be scrutinized to determine the number of components.

In previous work (5-9), this problem has been dealt with by analyzing the data using an increasingly larger number of components. The deviations between the observed and the calculated data are calculated and compared with the experimental error (thereby serving as the goodness-of-fit criterion). Usually the minimum number of components that can satisfactorily fit the observed data is taken to be the answer. It is recognized that if the data can be interpreted with (n) components, it can also be fitted to (n+1) components.

In this work, we propose to use factor analysis (10-13) as an alternative method whereby the number of components in a multicomponent polymer may be determined.

Factor analysis

Factor analysis (also known as principal component analysis) is a chemometric technique whereby analytical data may be represented in a small and manageable number of dimensions so as to observe groupings of objects, outliers, etc. which define the structure of a data set (11).

In NMR analysis of copolymers we typically deal with triads, and sometimes tetrads, pentads, and higher n-ads. For homopolymer tacticity we usually treat tetrads, pentads, or hexads. Without loss of generality, we shall confine our analysis to copolymer triads in this work. For ethylene/propylene copolymers, there are six triads (n=6): PPP, PPE, EPE, PEP, EEP, and EEE. For a series of m polymer mixtures, a data matrix [D] can be written with the dimension n × m.

A covariance matrix [C] is defined as follows:

$$[C] = [D]^t [D] \qquad (1)$$

The rank of the covariance matrix may be determined by solving the following eigenvalue problem:

$$[C] [E] = [E] [\lambda] \qquad (2)$$

where [E] is the matrix of eigenvectors and [λ] the diagonal matrix of eigenvalues. In the absence of noise and experimental error the number of non-zero eigenvalues is the number of independent components (factors). In the best cases involving experimental data there is a clear delineation between two groups of eigenvalues having large and small

magnitudes representing the pure and noise/error components, respectively. Alternatively, we may use Malinowski's indicator function (IND) (10):

$$IND = \frac{1}{(m - p)^2} \left[\frac{\sum_{j=p+1}^{m} \lambda_j}{n(m - p)} \right]^{1/2} \tag{3}$$

where p is the number of components. IND reaches a minimum when the appropriate number of factors is used. It is important to recognize that factor analysis determines the number of linearly independent components in a series of mixtures. As such the analysis can only detect compositional variation as reflected within the data set. For example, if the ratio of two species remains constant within the data set, even if their absolute concentration changes, only a single pure component will be associated with the pair.

Two examples of this treatment are given below. Both involve heterogeneous Ziegler-Natta catalysts which are known (14) to contain more than one active catalytic site.

Example 1

Xu, et al. (15) pre-polymerized a small amount of 1-octene using the $TiCl_3/Al(C_2H_5)_3$ catalyst and then proceeded to copolymerize ethylene and propylene with the same catalyst. They reported the triad data of copolymers made at several pre-polymerization times. The original data are shown in Table 1. A prior data analysis (7) has indicated that the polymer contains more than one component. This data set appears to be a good test case for factor analysis.

The [D] matrix can be readily set up:

$$[D] = \begin{bmatrix} 21.0 & 28.0 & 28.9 \\ 9.7 & 8.9 & 9.1 \\ 5.5 & 2.2 & 2.6 \\ 3.7 & 1.9 & 1.2 \\ 7.8 & 9.3 & 9.0 \\ 19.7 & 23.7 & 23.0 \end{bmatrix}$$

Each polymer sample is represented by a single column with rows within the column corresponding to concentrations of different triad species within the sample. Note that the intensities are not normalized. Calculation by Equations 1 and 2 can be made by the factor analysis routines described previously (12). The eigenvalues are plotted in Figure 1. The figure indicates that two components are appropriate for this system.

Another way to study the data is to regenerate the mixture data using the original data matrix coupled with eigenvectors associated with pure components. This approach essentially takes advantage of information present in all samples to improve the quality of individual measurements. The calculated data can then be compared with the observed data. The correct number of pure components will reconstruct "improved" estimates of the data associated with individual samples. Although use of eigenvectors beyond the actual number of pure components yields progressively smaller differences between the original and reconstructed datasets, the differences between using the actual number of pure components and excessive components are not statistically significant. These are given in Table 1. The two-component case does indeed produce smaller mean deviations

Table 1. NMR (unnormalized) triad data[a,b] for ethylene/propylene copolymers showing the effect of pre-polymerization with 1-octene.

Sample	pre-polym. time(h)		Triad Intensities					
			PPP	PPE	EPE	PEP	EEP	EEE
1	0.0	obsd.	21.0	9.7	5.5	3.7	7.8	19.7
		2-comp.	21.0	9.7	5.5	3.7	7.8	19.7
		1-comp.	22.6	7.9	2.8	1.9	7.5	19.2
2	0.5	obsd.	28.0	8.9	2.2	1.9	9.3	23.7
		2-comp.	28.3	9.1	2.6	1.6	9.1	23.3
		1-comp.	27.7	9.7	3.5	2.3	9.2	23.5
5	2.0	obsd.	28.9	9.1	2.6	1.2	9.0	23.0
		2-comp.	28.6	8.9	2.3	1.4	9.2	23.4
		1-comp.	27.9	9.7	3.5	2.3	9.3	23.6

[a]   Observed data from ref. 15.
[b]   For one-component analysis, SSQ = 23.3, mean deviation = 0.9.
For two-component analysis, SSQ = 0.9, mean deviation = 0.2.
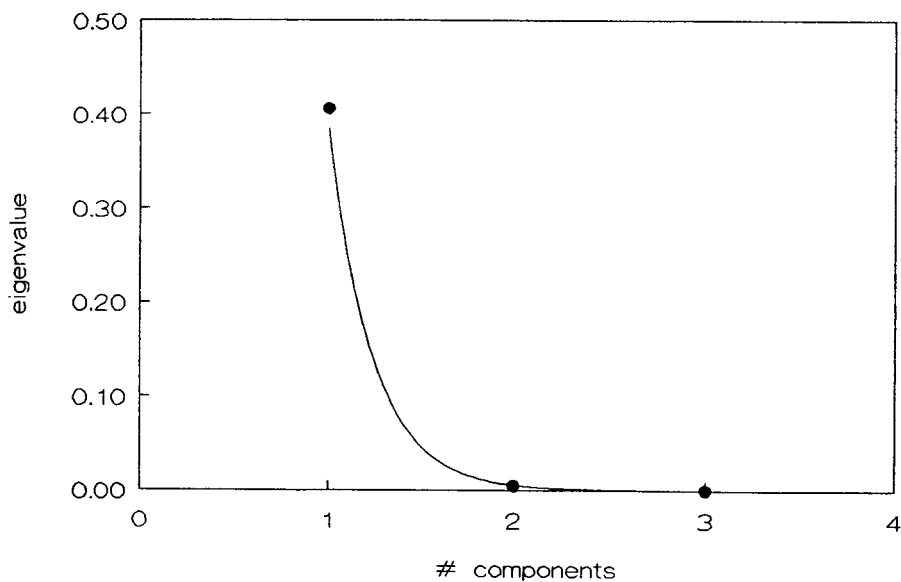(SSQ = sum of the squared deviations.)

Figure 1. Plot of eigenvalue versus component number for example 1.

than the one-component case (0.2% versus 0.9%). If the error in the data is on the order of 0.5%, it should be possible to distinguish the two-component case from the one-component case.

Example 2

Terano and Ishii (16) reported NMR triad data of several samples of ethylene/propylene copolymers made with the $TiCl_4$/ethyl benzoate/$MgCl_2$ catalyst subjected to mechanical grinding at various times (Table 2). The data have been previously analyzed and shown (7) to contain at least two components. The [D] matrix is:

$$[D] = \begin{bmatrix} 2 & 17 & 16 & 19 & 18 \\ 11 & 23 & 21 & 19 & 17 \\ 18 & 10 & 11 & 10 & 12 \\ 5 & 10 & 12 & 10 & 11 \\ 15 & 15 & 16 & 21 & 17 \\ 49 & 25 & 24 & 21 & 25 \end{bmatrix}$$

Computation is carried out as before and the eigenvalues are displayed in Figure 2. For this data set, the IND values are also calculated and displayed in Figure 3. It is clear that the data are best described by either 2 or 3 components. The IND values reach a minimum at p = 3.

We can again verify the number of components by regenerating the (theoretical) data matrix and assuming a varying number of pure components. The calculated data for one, two, and three eigenvalues are summarized in Table 2. It is important to note that the mean deviations between the observed and the reconstructed data are 4.0%, 0.9%, and 0.5%, corresponding to one-component, two-component, and three-component cases respectively. Thus, how well we can resolve the data into two or three components depends, in part, on the precision and the accuracy of the NMR data. In this case, if the experimental error of the data is 0.5% or less, then we can justifiably use three components to carry out further analysis. If the data error is around 1%, then the data, on their own merit, do not justify the use of three components.

Comments

It is perhaps not surprising that the ability to ascertain the number of components depends on the quality and the quantity of data available. Thus, the smaller the experimental error, the more component(s) we can discern from data analysis. Moreover, the more data there are, the more information we can obtain through analysis. As a case in point, we can probably treat example 2 (with a larger [D] matrix) as three-component copolymers, whereas the more limited data of example 1 permit its treatment only as two-component copolymers.

For convenience, the methodologies described herein are summarized in a scheme (Figure 4). Factor analysis has the advantages that equal weight is placed on each data point and that the evaluation of the data is relatively objective. Although only copolymer triad data are illustrated in this work, this analysis can be readily extended to higher order sequences (e.g., tetrads, pentads, and hexads) in copolymers and tacticity data of homopolymers.
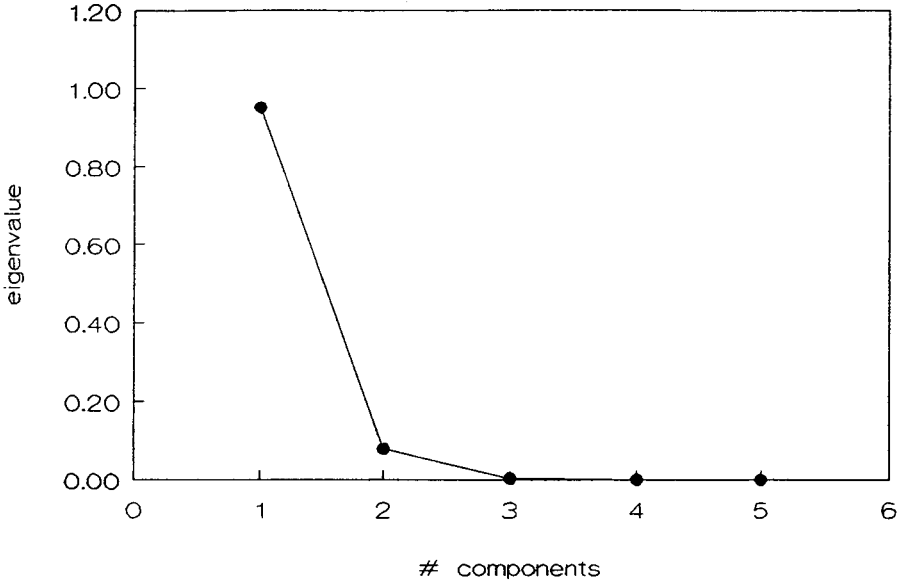
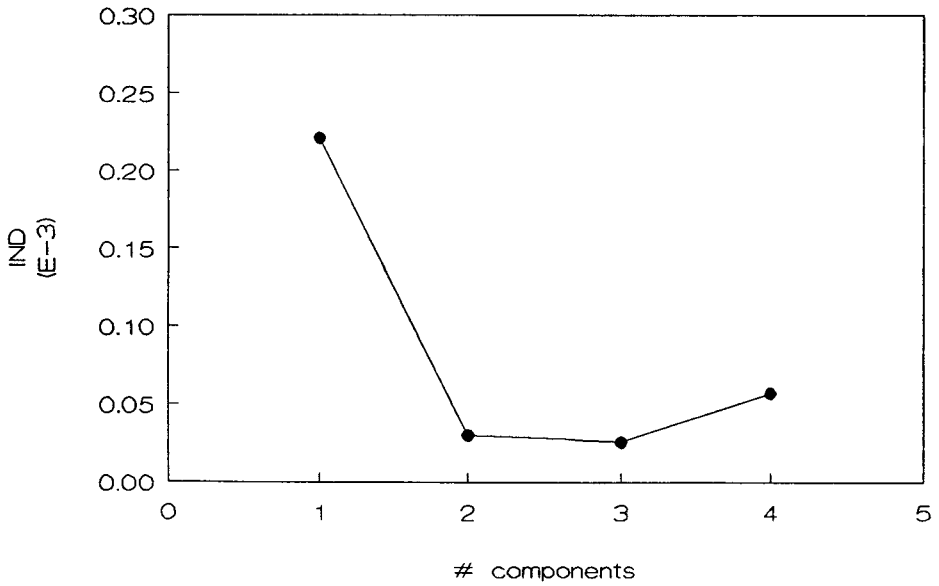Figure 2. Plot of eigenvalue versus component number for example 2.

Figure 3. Plot of IND value versus component number for example 2.

Table 2. NMR triad data[a,b] for ethylene/propylene copolymers showing the effect of catalyst grinding times.

Sample	grinding time(h)		Triad Intensities PPP	PPE	EPE	PEP	EEP	EEE
A_0	0	obsd.	2.0	11.0	18.0	5.0	15.0	49.0
		3-comp.	2.1	10.9	18.1	5.1	14.9	49.0
		2-comp.	2.1	11.1	18.1	5.2	14.6	49.1
		1-comp.	16.0	20.6	14.3	10.8	19.2	34.0
A_5	5	obsd.	17.0	23.0	10.0	10.0	15.0	25.0
		3-comp.	16.7	22.8	10.5	11.1	14.9	24.8
		2-comp.	17.5	20.1	11.0	10.8	17.5	24.4
		1-comp.	13.5	17.4	12.1	9.1	16.2	28.7
A_{10}	10	obsd.	16.0	21.0	11.0	12.0	16.0	24.0
		3-comp.	16.7	21.1	10.6	10.7	15.9	24.3
		2-comp.	17.1	19.7	10.8	10.6	17.2	24.1
		1-comp.	13.3	17.1	11.9	9.0	16.0	28.3
A_{15}	15	obsd.	19.0	19.0	10.0	10.0	21.0	21.0
		3-comp.	19.7	18.3	10.6	10.9	20.1	21.1
		2-comp.	18.9	20.8	10.1	11.2	17.6	21.5
		1-comp.	13.1	16.8	11.7	8.8	15.7	27.8
A_{30}	30	obsd.	18.0	17.0	12.0	11.0	17.0	25.0
		3-comp.	16.8	18.0	11.3	10.1	18.3	24.8
		2-comp.	16.4	19.3	11.1	10.3	17.0	25.0
		1-comp.	13.3	17.2	11.9	9.0	16.0	28.3

[a] Observed data from ref.16.

[b] For one-component analysis, SSQ = 848.6, mean deviation = 4.0.
For two-component analysis, SSQ = 50.0, mean deviation = 0.9.
For three-component analysis, SSQ = 12.5, mean deviation = 0.5.
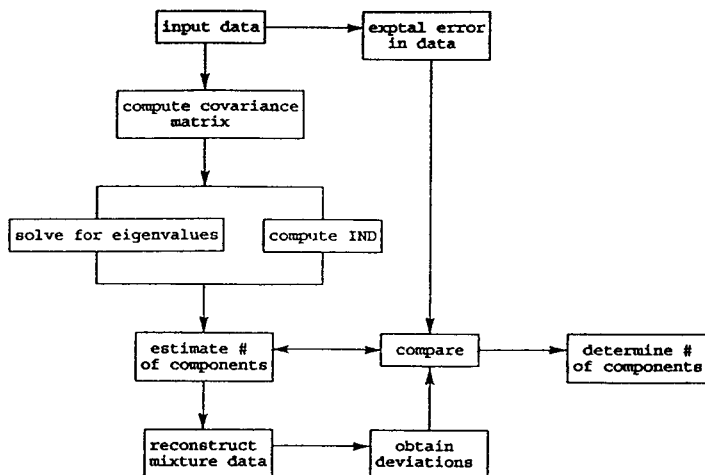(SSQ = sum of the squared deviations.)

Figure 4. Factor analysis scheme used for the determination of the number of components.

References

1. Coleman BD, Fox TG (1963) *J Chem Phys* **38**:1065.
2. Zambelli A, Locatelli P, Provasoli A, Ferro DR (1980) *Macromolecules* **13**: 267.
3. Inoue Y, Itabashi Y, Chujo R, Doi Y (1984) *Polymer* **25**:1640.
4. Ross JF (1988). In Quirk RP (ed) *Transition Metal Catalyzed Polymerizations.* Cambridge Univ. Press Cambridge, p. 799.
5. Cheng HN (1988) *J Appl Polym Sci* **35**:1639.
6. Cheng HN (1989) *ACS Symp Ser* **404**:174.
7. Cheng HN (1993) *Makromol Chem Theory Simul* **2**:901.
8. Cheng HN, Kakugo M (1991) *Macromolecules* **24**:1724.
9. Cheng HN (1992) *Makromol Chem Theory Simul* **1**:415.
10. Malinowski ER, Howery DG (1980) *Factor Analysis in Chemistry.* Wiley, New York.
11. Massart DL, Vandeginste BGM, Deming SN, Michotte Y, Kaufman I (1988) *Chemometrics: A Textbook.* Elsevier, Amsterdam.
12. Gillette PC, Lando JB, Koenig JL (1985). In Ferraro JR, Basile LJ (eds), *Fourier Transform Infrared Spectroscopy*, Academic Press, Orlando, vol. 4, p. 1.
13. Gillette PC, Lando JB, Koenig JL (1983) *Anal Chem* **55**:630.
14. For example, (a) Kissen YV (1985) *Isospecific Polymerization of Olefins with Heterogeneous Ziegler-Natta Catalysts*, Springer-Verlag, New York, chap. 4; (b) Tait PJT (1989). In Allen G, Bevington JC (eds) *Comprehensive Polymer Science*, Pergamon, Oxford. (c) Chien JCW (1992) *ACS Symp Ser* **496**:27.
15. Xu Z, Zhu Q, Feng L, Yang S (1990) *Makromol Chem Rapid Comm* **11**:79.
16. Terano M, Ishii K (1991) *Makromol Chem Rapid Comm* **11**:439; also *Polym Comm* **32**: 314.