



ACADEMIC
PRESS

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Magnetic Resonance 164 (2003) 10–18

JMR

Journal of
Magnetic Resonance

www.elsevier.com/locate/jmr

An efficient peak assignment algorithm for two-dimensional NMR correlation spectra of framework structures[☆]

Darren H. Brouwer*

Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, BC, Canada V6T 1Z1

Received 23 September 2002; revised 26 May 2003

Abstract

An algorithm is described for efficiently assigning the resonances in NMR spectra to the inequivalent atoms in the structure under study based on the information in two-dimensional NMR correlation experiments and the ‘connectivities’ known from the structure. The algorithm, which is based on basic graph theory concepts, finds *all* possible assignments sets which are consistent with the experimentally observed correlations and known connectivities in a very efficient manner. It is designed to deal with less than ideal experimental data in which there may be overlapping peaks and uncertainty about the presence or absence of correlation peaks. The algorithm was primarily developed for assigning the peaks in the high-resolution solid-state ²⁹Si MAS NMR spectra of highly siliceous zeolites based on two-dimensional ²⁹Si INADEQUATE spectra and is described using the zeolites ZSM-12 and ZSM-5 as working examples. Peak assignment for zeolite frameworks is particularly challenging since there is often little or no information to distinguish peaks from one another such as characteristic chemical shifts, relative intensities, or different relaxation times. The algorithm may be a useful tool for easily, reliably, and efficiently working out peak assignments from other types of correlation experiments on other types of systems and further examples are provided in the Supplementary material.

© 2003 Elsevier Inc. All rights reserved.

1. Introduction

Multi-dimensional NMR correlation experiments provide a wealth of structural information as they probe the ‘connectivities’ between atoms that exist within the system under study. These powerful experiments can utilize heteronuclear interactions for correlating the spectrum of one type of nucleus to the spectrum of another type or homonuclear interactions for correlating nuclei of the same type. All of these experiments yield multi-dimensional correlation spectra which indicate which peaks are ‘connected’ to each other via the particular interaction exploited by the NMR experiment. This paper describes an efficient algorithm for relating such experimental correlation spectra to the known connectivities that are present in the structure under

study in order to assign the individual resonances in the NMR spectrum to the inequivalent sites in the structure.

The assignment of the peaks in ¹H, ¹³C, and ¹⁵N spectra of proteins from multi-dimensional NMR experiments is a crucial step in the determination of their three-dimensional structures and various approaches have been developed [1]. For organic molecules, the connectivity information available from two-dimensional ¹H and ¹³C experiments can be used for elucidating the structure of unknown organic compounds [2]. These approaches to peak assignment take advantage of the great deal of additional structural information available from characteristic chemical shifts and coupling patterns which is incorporated with the information provided by multi-dimensional correlation experiments. For example, the chemical shifts in ¹H or ¹³C spectra of organic compounds are characteristic of the local environments of the ¹H or ¹³C nuclei and can be used to identify specific peaks as methyl, methylene, aromatic, carbonyl, etc. With this knowledge, it is relatively straightforward to work out the peak assignments from the information provided by correlation experi-

[☆] Supplementary data associated with this article can be found on ScienceDirect.

* Fax: +604-822-2847.

E-mail address: dbrouwer@chem.ubc.ca.

ments. However, as the number of inequivalent atoms in the structure increases and in situations where there is less additional information available from characteristic chemical shifts, relative intensities, differences in relaxation times, or from other experiments, it becomes increasingly difficult to work out the peak assignments from correlation experiments and to be confident that all possible sets of peak assignments have been tested.

The assignment of peaks in the high-resolution solid-state ^{29}Si magic angle spinning (MAS) NMR spectra of highly siliceous zeolites to the inequivalent Si atoms in the zeolite framework is a crucial first step in the determination of the three-dimensional structures of zeolite host–guest complexes involving ions [3] or small organic molecules [4] from internuclear distance information obtained from cross polarization or REDOR experiments. Solid state NMR is an important alternative and complementary technique to diffraction methods for the determination of zeolite host–guest structures due to the extremely small size of most zeolite crystals which are usually on the order of micrometers. The two-dimensional ^{29}Si INADEQUATE experiment [5] which probes the indirect ^{29}Si –O– ^{29}Si spin–spin couplings has successfully been applied to a number of different highly siliceous zeolites at natural abundance [6–8]. The resulting peak assignments have been used to develop correlations between ^{29}Si chemical shifts and information about local environments available from a small number of reliable X-ray diffraction structures [9].

The assignment of the peaks in the ^{29}Si spectra of highly siliceous zeolites is particularly challenging. In many zeolite structures, the site occupancies of the Si atoms are often identical such that the relative intensities of the peaks in the ^{29}Si spectrum do not provide any information about the identity of any of the peaks. Also, since all of the Si atoms are found in tetrahedral coordination environments, the ^{29}Si chemical shifts usually provide no information about the identity of the peaks. Often, the only source of information to work with is the ^{29}Si – ^{29}Si homonuclear correlations observed in the two-dimensional INADEQUATE spectrum. Without any knowledge of the identity of any of the peaks, working out the peak assignments from only the observed correlations can be a very challenging task and is well-suited for analysis by computer.

This paper describes an algorithm, based on basic graph theory concepts, which efficiently finds *all* possible peak assignments for which there is agreement between the experimentally observed correlations and the known connectivities in the structure. The algorithm is designed to deal with less than ideal experimental data in which there may exist overlapping peaks and uncertainty regarding the presence or absence of correlation peaks. The algorithm is described in detail using the ^{29}Si INADEQUATE spectra of a number of zeolites as working examples. Although this algorithm was developed

for assigning peaks from the ^{29}Si INADEQUATE spectra of zeolites, it may be a useful tool for easily, reliably, and efficiently working out peak assignments from other types of correlation experiments (solution or solid state) on other types of samples, especially in situations for which there exists little information about the identity of the peaks other than the correlation data.

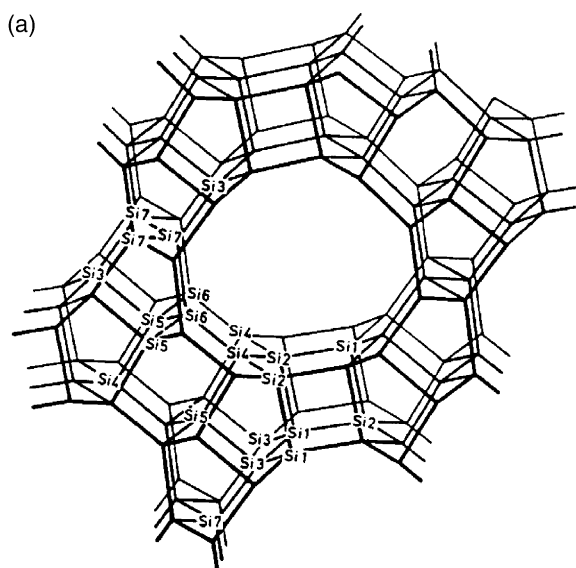
2. Description of the algorithm

The peak assignment algorithm uses some basic graph theory concepts which are briefly described here. A ‘graph’ $G(V, E)$ consists of a set of points called ‘vertices’ V and a set of lines called ‘edges’ E which connect these points. An ‘adjacency matrix’ $A(G)$ is a mathematical matrix representation of a graph in which each i, j element is the number of edges between vertices i and j . These basic graph theory concepts can be used to represent both the connectivity pattern in the structure under study and the correlation pattern observed in the two-dimensional NMR experiment.

Consider the example of the zeolite ZSM-12 [6,10]. To describe the framework topology of ZSM-12 (Fig. 1a) as a graph, the set of vertices corresponds to the seven inequivalent Si sites and the set of edges corresponds to the Si–O–Si bonds which exist between them (Fig. 1b). A graph representation of the ZSM-12 zeolite framework topology is shown in Fig. 1c and the corresponding adjacency matrix is given in Fig. 1d. Note that the connectivities between Si7 are not used since ‘self-connectivities’ do not appear in the experimental INADEQUATE spectrum.

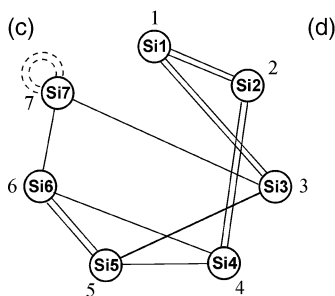
The two-dimensional ^{29}Si INADEQUATE spectrum (Fig. 2a) can also be represented by a graph. The set of vertices corresponds to the labels given to the peaks and the set of edges corresponds to the observed correlations between the peaks (Fig. 2b). Note that in this example the more intense peaks were assigned as two correlations while the weaker peaks were assigned as one correlation. A graph representation of the ^{29}Si INADEQUATE spectrum of the ZSM-12 zeolite is shown in Fig. 2c and the corresponding adjacency matrix is given in Fig. 2d.

A proposed assignment $A \rightarrow P$, where A is a set of atoms and P is a set of peaks, can be tested by comparing the graphs $G(A, C_A)$ and $G(P, C_P)$ where C_A is the list of connectivities between the atoms and C_P is the list of observed correlations between the peaks. An assignment is said to be ‘good’ if the graphs are equivalent, or in mathematical terms, the adjacency matrices are equal. For example, the assignment $\{\text{Si2, Si5, Si7, Si1, Si6, Si4, Si3}\} \rightarrow \{\mathbf{A, B, C, D, E, F, G}\}$ is ‘good’ because the two graphs and their corresponding adjacency matrices are equivalent, as shown in Fig. 3a. This test can also be applied for assessing the assignment of a subset of the peaks: the assignment $\{\text{Si4, Si5, Si1, Si3}\} \rightarrow \{\mathbf{F, B, D, G}\}$



(b) Set of connectivities between atoms (C_A):

Si1 – Si2 (2)	Si3 – Si5 (1)	Si5 – Si6 (2)
Si1 – Si3 (2)	Si4 – Si5 (1)	Si6 – Si7 (1)
Si2 – Si4 (2)	Si4 – Si6 (1)	Si7 – Si7 (2)*



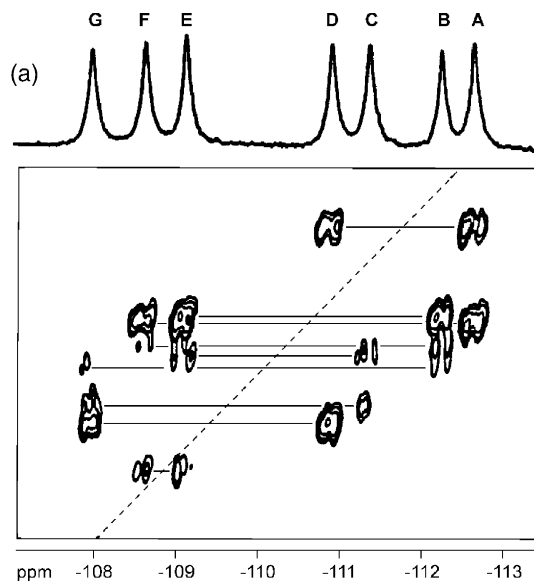
(d)

$$\begin{bmatrix} 0 & 2 & 2 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 2 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Fig. 1. The structure of the zeolite ZSM-12: (a) illustration of the framework structure and Si atom numbering (reprinted in part from [6]. Copyright (1990) American Chemical Society); (b) set of connectivities between the Si atoms; (c) a graph representation of the ZSM-12 framework; and (d) corresponding adjacency matrix. *Note that ‘self-connectivities’ between Si7 are not used as they do not appear in the INADEQUATE spectrum.

is ‘good’ since the two graphs and their corresponding adjacency matrices are equivalent, as shown in Fig. 3b. Note that the graphs and corresponding adjacency matrices are a function of the order in which the atoms or peaks are listed in A or P .

Using this approach to testing proposed assignments, it is possible to find *all* of the peak assignments which give agreement between the observed correlations and known connectivities. The ‘brute-force’ approach to this task is to propose and test each and every possible combination of peak assignments. The total number of proposed assignments that needs to be tested is $N!$ where N is the number of peaks or inequivalent atoms. In this example of ZSM-12, the total number of assignment sets is $7! = 5040$ which is a reasonable number to test.



(b) Set of correlations between peaks (C_P):

A – D (2)	B – F (1)	D – G (2)
A – F (2)	B – G (1)	E – F (1)
B – E (2)	C – G (1)	

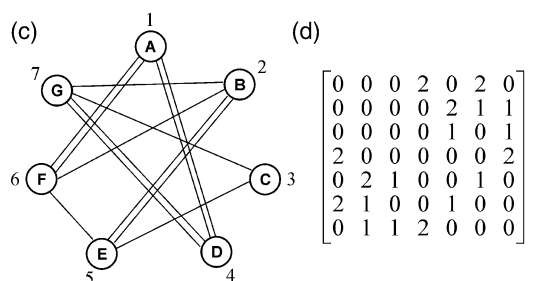


Fig. 2. (a) Two-dimensional ^{29}Si INADEQUATE spectrum of the zeolite ZSM-12 (reprinted in part from [6]. Copyright (1990) American Chemical Society); (b) set of observed correlations between the peaks; (c) a graph representation of the INADEQUATE spectrum; and (d) corresponding adjacency matrix.

However as N gets larger, the number of possible assignments becomes too many to test in a reasonable amount of time. For example, solving the INADEQUATE spectrum in this manner for the monoclinic phase of ZSM-5 (see later) in which there are 24 inequivalent Si atoms is just not possible as $24! = 6.2 \times 10^{23}$. Therefore a strategy is required to greatly reduce the number of combinations tested, making even the situation with $N = 24$ possible to solve.

In order to reduce the number of combinations tested, yet still ensuring that *all* of the assignments which are consistent with the data are found, a ‘tree search algorithm’ is employed. In this approach, the peaks are assigned and tested in a step-wise and ‘branching’ manner in which only the good assignment combinations are passed on from one step to the next. To illustrate, peak {A} is first assigned to {Si1}, {Si2}, {Si3},

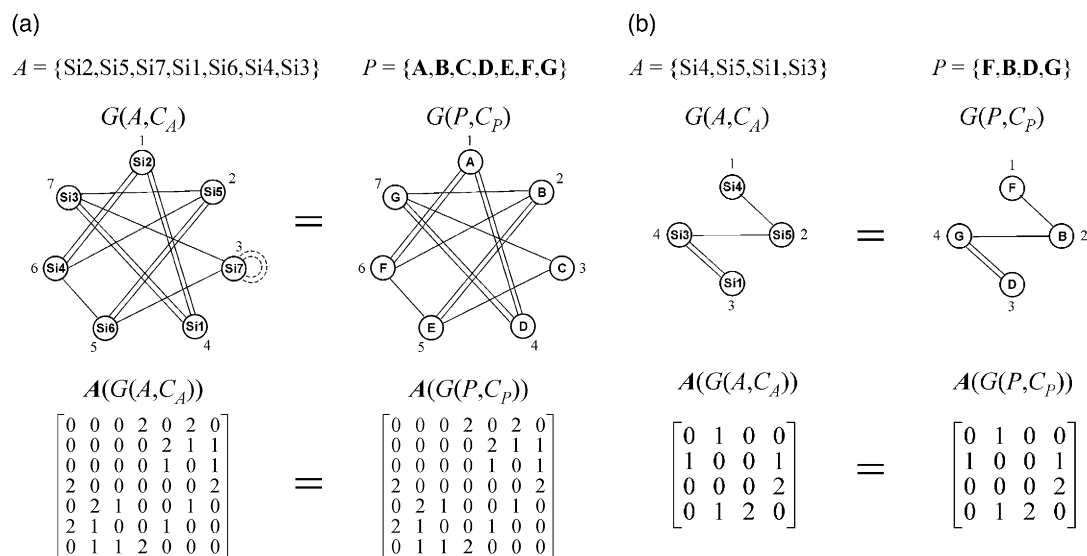


Fig. 3. The assignments $A \rightarrow P$ are ‘good’ since the graphs and corresponding adjacency matrices are equal: (a) a full assignment in which all of the peaks are tested, (b) a partial assignment in which a subset of the peaks are tested. Note that graphs and adjacency matrices are a function of the order in which the atoms and peaks are listed in A and P .

$\{\text{Si}4\}$, $\{\text{Si}5\}$, $\{\text{Si}6\}$, and $\{\text{Si}7\}$ and tested as described above. The graphs consist of a single vertex without edges and the adjacency matrices of a single element of zero. All of these assignments give graphs equal to $G(\{\mathbf{A}\}, C_P)$ and pass the test. In the next step, peaks $\{\mathbf{A}, \mathbf{B}\}$ are tested against assignment combinations obtained by ‘branching’ from the ‘good’ assignments in the previous step. For example, $\{\text{Si}1\}$ branches to $\{\text{Si}1, \text{Si}2\}$, $\{\text{Si}1, \text{Si}3\}$, $\{\text{Si}1, \text{Si}4\}$, $\{\text{Si}1, \text{Si}5\}$, $\{\text{Si}1, \text{Si}6\}$, and $\{\text{Si}1, \text{Si}7\}$, while $\{\text{Si}2\}$ branches to $\{\text{Si}2, \text{Si}1\}$, $\{\text{Si}2, \text{Si}3\}$, $\{\text{Si}2, \text{Si}4\}$, $\{\text{Si}2, \text{Si}5\}$, $\{\text{Si}2, \text{Si}6\}$, and $\{\text{Si}2, \text{Si}7\}$, and so on for $\{\text{Si}3\}$, $\{\text{Si}4\}$, $\{\text{Si}5\}$, $\{\text{Si}6\}$, and $\{\text{Si}7\}$. Out of this set of proposed assignments, only the assignment combinations which have graphs equivalent to $G(\{\mathbf{A}, \mathbf{B}\}, C_P)$ are passed on to the next step in which these ‘good’ assignments are branched again and the graphs compared to $G(\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}, C_P)$. These steps are repeated until all of the peaks are assigned. The end result is a list of *all* of the possible sets of peak assignments which give full agreement between the experimentally observed correlations and the known connectivities between the atoms in the structure.

The application of this algorithm to the ZSM-12 example is shown in Fig. 4 in which one ‘branch’ of the search for fully consistent peak assignments is illustrated. The algorithm searches the entire solution space and finds the one correct set of peak assignments by testing only 294 assignment combinations in 0.38 s. This is a very large gain in efficiency compared to the ‘brute-force’ approach to searching the entire solution space which tests $7! = 5040$ assignment combinations in 8.7 s.

The above example of ZSM-12, in which all seven peaks are fully resolved and all of the correlations between them are observed, is ideal for illustrating

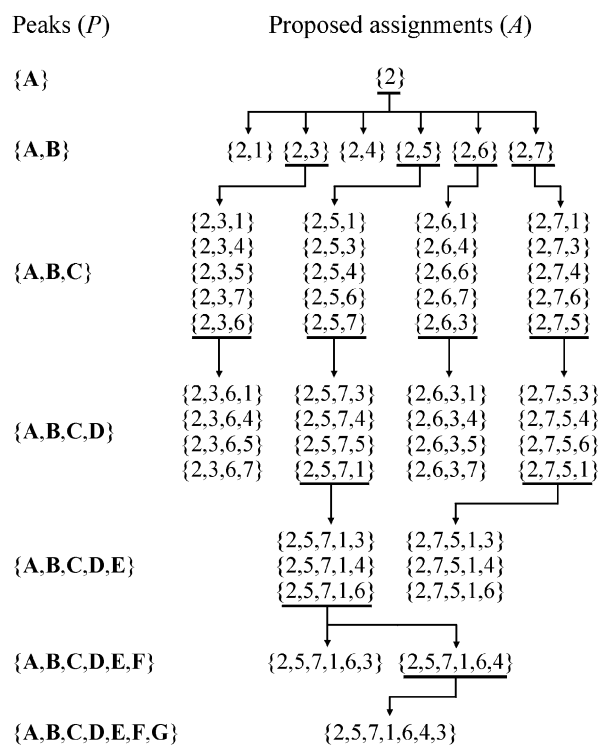


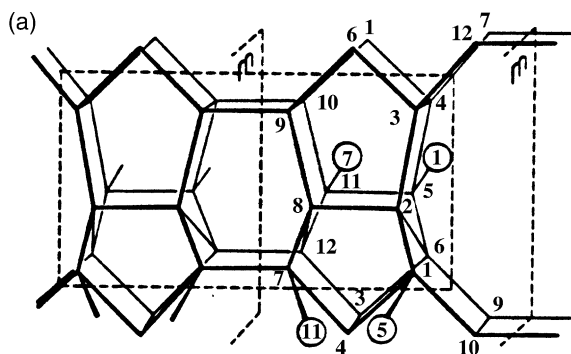
Fig. 4. One branch of the ‘tree search algorithm’ for ZSM-12 which searches for all ‘good’ assignments $A \rightarrow P$ in a step-wise and branching manner. The underlined assignments are found to be ‘good’ since the graphs $G(A, C_A)$ and $G(P, C_P)$ are equivalent. These assignment combinations are passed on to the next step in which they are ‘branched’ and tested again. The steps are repeated until all seven of the peaks are assigned. There is only one assignment combination which is in full agreement with the experimental data. The numbers refer to the Si atoms.

the inner workings of this peak assignment algorithm. However, in reality, the experimental data are rarely this ideal. The most common problems for ^{29}Si

INADEQUATE spectra of zeolites are peak overlap and uncertainty about the existence or intensity of some of the correlations. Consequently, the algorithm was modified to account for less than ideal experimental data.

These modifications are illustrated with another example, the ^{29}Si INADEQUATE spectrum of the high temperature orthorhombic phase of the zeolite ZSM-5 [7] in which there are 12 inequivalent Si atoms. The zeolite framework and Si atom numbering scheme are depicted in Fig. 5a. The connectivities between the Si atoms (C_A) are listed in Fig. 5b. Note that the four different self connectivities are not used in this case because, again, they are not observed in the INADEQUATE experiment.

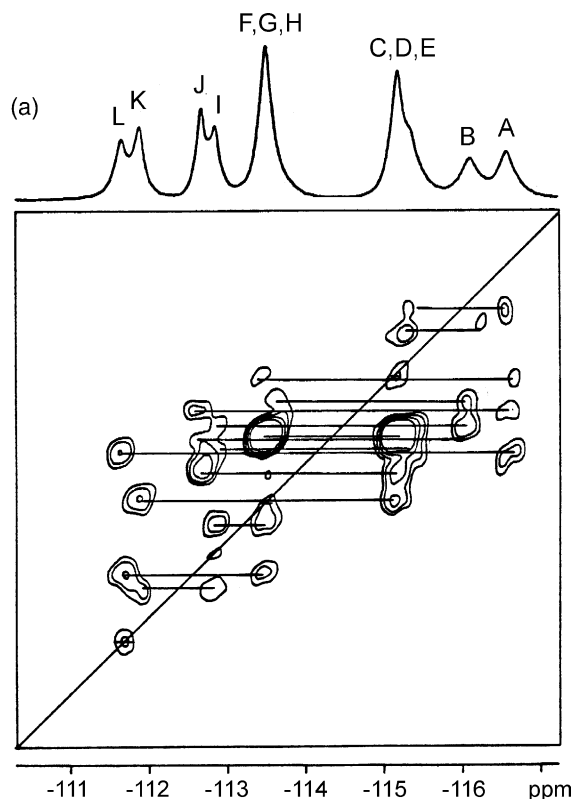
The two-dimensional ^{29}Si INADEQUATE spectrum is presented in Fig. 6a. Only six of the 12 peaks are fully resolved as there are two groups of overlapping peaks each consisting of three peaks based on the relative intensities of the peaks. The correlations between the peaks (C_P) are listed in Fig. 6b and this list includes the uncertainty and ambiguity of some of the correlations arising from the overlap of the peaks. For example, peak A is definitely correlated to peaks J and L but is correlated to only one of peaks C, D, E, and to only one of peaks F, G, H. As another example, there is either one or two correlations between peaks C, D, E and I, J and there is an uncertain number of correlations between the overlapping groups of peaks C, D, E and F, G, H. Also,



(b) Set of connectivities between the atoms (C_A):

Si1 – Si2	Si3 – Si4	Si6 – Si9	Si10 – Si11
Si1 – Si4	Si3 – Si6	Si7 – Si8	Si11 – Si12
Si1 – Si5	Si3 – Si12	Si7 – Si11	Si7 – Si7*
Si1 – Si10	Si4 – Si5	Si8 – Si9	Si9 – Si9*
Si2 – Si3	Si4 – Si7	Si8 – Si12	Si10 – Si10*
Si2 – Si6	Si5 – Si6	Si9 – Si10	Si12 – Si12*
Si2 – Si8	Si5 – Si11		

Fig. 5. The structure of the high temperature orthorhombic phase of the zeolite ZSM-5: (a) illustration of the framework structure and Si atom numbering (reprinted in part from [7]. Copyright (1990) American Chemical Society); (b) set of connectivities between the Si atoms. *Note that 'self-connectivities' between Si7, Si9, Si10, and Si12 are not used as they do not appear in the INADEQUATE spectrum.



(b) Set of correlations between the peaks (C_P):

A – C,D,E (1)	C,D,E – C,D,E (?)
A – F,G,H (1)	C,D,E – F,G,H (?)
A – J (1)	C,D,E – I,J (1,2)
A – L (1)	C,D,E – K (1)
B – C,D,E (1)	F,G,H – F,G,H (?)
B – F,G,H (1)	F,G,H – I (1)
B – I (1)	F,G,H – L (1)
B – J (1)	I – K (1)
I – J (0,1)	K – L (1)

Fig. 6. (a) Two-dimensional ^{29}Si INADEQUATE of the high temperature orthorhombic phase of the zeolite ZSM-5 (reprinted in part from [7]. Copyright (1990) American Chemical Society); (b) set of the observed correlations between the peaks.

the small peak on the diagonal near peaks I and J may be an artifact or may actually be a correlation between I and J.

The adjacency matrix that is constructed from this list of correlations includes the uncertainty and ambiguity in the correlations by assigning the value '0, 1' to any edge between two vertices for which the connection may be uncertain or ambiguous, indicating that either zero or one correlation may be present. For example, the entry 'A–C, D, E (1)' in the list of correlations means that the A–C, A–D, and A–E elements in the adjacency matrix are each assigned the value '0, 1.' The complete adjacency matrix constructed from the list of correlations in Fig. 6b is shown in Fig. 7.

	A	B	C	D	E	F	G	H	I	J	K	L
A	0	0	0,1	0,1	0,1	0,1	0,1	0,1	0	1	0	1
B	0	0	0,1	0,1	0,1	0,1	0,1	0,1	1	1	0	0
C	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0
D	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0
E	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0
F	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0	0	0,1
G	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0	0	0,1
H	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0	0	0,1
I	0	1	0,1	0,1	0,1	0,1	0,1	0,1	0	0,1	1	0
J	1	1	0,1	0,1	0,1	0	0	0	0,1	0	0	0
K	0	0	0,1	0,1	0,1	0	0	0	1	0	0	1
L	1	0	0	0	0	0,1	0,1	0,1	0	0	1	0

Fig. 7. Adjacency matrix constructed from the list of observed correlations in the ^{29}Si INADEQUATE spectrum of the high temperature orthorhombic phase of the zeolite ZSM-5 (see Fig. 6). An entry of ‘0, 1’ indicates that there may exist zero or one correlation(s) between the peaks.

When a proposed assignment $A \rightarrow P$ is tested by comparing the adjacency matrices $A(G(A, C_A))$ and $A(G(P, C_P))$, the condition for a ‘good’ assignment is no longer a rigid condition in which the two adjacency matrices must be exactly equal, but is rather a more flexible condition in which the uncertainty and ambiguity in the correlations between the peaks, denoted by the ‘0, 1’ elements, is taken into account. For example, the assignment $\{\text{Si8, Si4, Si2, Si3}\} \rightarrow \{\text{A, B, C, D}\}$ is ‘good’ since the corresponding adjacency matrices are consistent and in agreement with each other since a number of the elements in the $A(G(A, C_A))$ matrix can be ‘0’ or ‘1’, as illustrated in Fig. 8. With this modification, this peak assignment algorithm is quite flexible and can accommodate less than ideal experimental data in which there is peak overlap and uncertainty in the absence or presence of correlations.

The order in which the peaks are tested in the tree search algorithm impacts the efficiency of the algorithm in terms of the total number of peak assignment combinations tested. If the peaks for which more information is known are tested at the initial stages of the search, there will be fewer ‘branches’ of peak assignment combinations for the algorithm to search. An efficient

$$\begin{array}{l}
 A = \{\text{Si8, Si4, Si2, Si3}\} \qquad P = \{\text{A, B, C, D}\} \\
 \\
 \begin{array}{c}
 A(G(A, C_A)) \\
 \begin{bmatrix}
 0 & 0 & 1 & 0 \\
 0 & 0 & 0 & 1 \\
 1 & 0 & 0 & 1 \\
 0 & 1 & 1 & 0
 \end{bmatrix}
 \end{array}
 \qquad = \qquad
 \begin{array}{c}
 A(G(P, C_P)) \\
 \begin{bmatrix}
 0 & 0 & 0,1 & 0,1 \\
 0 & 0 & 0,1 & 0,1 \\
 0,1 & 0,1 & 0,1 & 0,1 \\
 0,1 & 0,1 & 0,1 & 0,1
 \end{bmatrix}
 \end{array}
 \end{array}$$

Fig. 8. For the high temperature orthorhombic phase of the zeolite ZSM-5, the assignment $A \rightarrow P$ is ‘good’ since the adjacency matrices are in agreement. Note the the ‘0, 1’ elements indicate that there may exist zero or one connection(s).

order in which to test the peaks can be obtained by ranking the peaks according to the number of correlations which are definitely absent or present (i.e., ranked by the number of ‘0’ and ‘1’ elements in a row or column of the adjacency matrix). In the example of the high temperature orthorhombic phase of ZSM-5, an efficient order for testing the peaks is $\{\text{K, L, J, A, B, I, F, G, H, C, D, E}\}$. Using this order for testing the peaks, the algorithm finds 6696 possible peak assignment combinations which give full adjacency matrices in agreement with each other in 73 s by testing only 4.6×10^5 out of $12! = 4.8 \times 10^8$ assignment combinations.

These assignments are then further tested to ensure that they are fully consistent with the list of correlations. For example, each assignment is tested to make sure that the Si atom assigned to peak A has one and only one connection to the Si atoms assigned to C, D, and E so that the condition ‘A–C, D, E (1)’ is satisfied. After testing each element in the list of correlations for each peak assignment combination found by the tree search algorithm, the number of ‘good’ assignment combinations is further reduced to 216 (in 146 s).

Lastly, the ‘good’ assignment combinations which are degenerate because of the peak overlap are grouped together. For example, there will be degenerate assignment combinations with all permutations of $\{\text{Si5, Si6, Si12}\}$ assigned to $\{\text{F, G, H}\}$ since there is nothing to distinguish between peaks F, G, and H. For the high temperature orthorhombic ZSM-5 example, the ‘good’ assignments are grouped into six non-degenerate peak assignment combinations which are fully consistent with the experimental data.

Since there exists an empirical correlation between the ^{29}Si chemical shifts and the mean Si–Si distances [9], it is possible to identify which peak assignments are the most reasonable. The six non-degenerate peak assignment combinations are listed in Table 1 with the degree of linear correlation (r^2) between the ^{29}Si chemical shifts and mean Si–Si distances from the single crystal XRD structure [11]. There is only one assignment which gives a high degree of correlation.

These modifications to this peak assignment algorithm give it a great deal of flexibility so that it can accommodate less than ideal experimental data in which there exists peak overlap and uncertainty and ambiguity about the presence or absence of correlations between the peaks, yet still ensuring that *all* possible assignment combinations which agree with the experimental correlation data are found in a very efficient manner. In fact, this peak assignment algorithm has been successfully applied to a fluoride-containing as-synthesized zeolite in which all of the correlations to one of the ^{29}Si peaks are completely absent due to a dynamic exchange process involving the fluoride ion [3].

As a final example, this algorithm is applied for the peak assignment of the ^{29}Si MAS NMR spectrum of the

Table 1
Possible assignments for the ^{29}Si MAS NMR spectrum of the orthorhombic phase of zeolite ZSM-5

A	B	{C, D, E}	{F, G, H}	I	J	K	L	r^{2a}
8	4	{2, 3, 11}	{12, 5, 6}	1	7	10	9	0.947
8	4	{5, 11, 12}	{2, 3, 6}	1	7	10	9	0.583
11	3	{4, 5, 8}	{1, 2, 7}	6	12	9	10	0.510
11	3	{2, 7, 8}	{1, 4, 5}	6	12	9	10	0.359
6	4	{3, 11, 12}	{2, 7, 8}	1	5	10	9	0.252
1	2	{4, 7, 8}	{5, 11, 12}	6	2	9	10	0.154

^a Assignments are ranked according to the degree of linear correlation between the ^{29}Si chemical shifts and the mean Si–Si distances.

room temperature monoclinic phase of the zeolite ZSM-5 [7,12]. This is a very difficult spectrum to assign since there are 24 inequivalent Si atoms, many of the peaks are overlapping, and there is no additional information outside of the ^{29}Si INADEQUATE spectrum to identify the peaks (except the empirical correlation between ^{29}Si chemical shifts and mean Si–Si distances). This challenging example clearly demonstrates the usefulness of this peak assignment algorithm.

The ^{29}Si INADEQUATE spectrum of the monoclinic phase of ZSM-5 is shown in Fig. 9. The list of connectivities between the Si atoms (C_A) can be found in [7] and

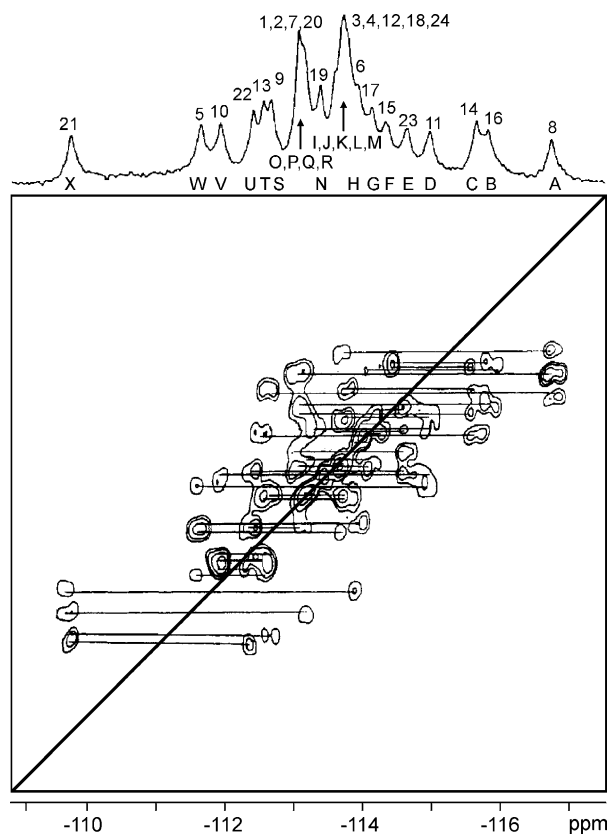


Fig. 9. Two-dimensional ^{29}Si INADEQUATE spectrum of the room temperature monoclinic phase of the zeolite ZSM-5 with the peak assignments found by the algorithm. The spectrum is reprinted in part from [7]. Copyright (1990) American Chemical Society.

is supplied in the Supplementary material. The list of correlations between the peaks (C_P) extracted from this spectrum is also provided in the Supplementary material. The peak assignment algorithm finds four non-degenerate assignment combinations which are fully consistent with the experimental data in 5.5 h by testing 5.8×10^6 assignment combinations (compared to $24! = 6.2 \times 10^{23}$). Only one of these assignments (shown in Fig. 9) gives a strong correlation between the ^{29}Si chemical shifts and Si–Si distances obtained from the single crystal XRD structure [13].

The amount of time required by the algorithm to solve a given peak assignment problem depends on a number of factors. The most obvious factor is the number of peaks or inequivalent atoms. Second, the amount of time will depend on how much information is available about the identity of any of the peaks (based on characteristic chemical shifts, relative intensities, relaxation times, etc.). The amount of time required by the algorithm also depends on the ‘quality’ of the correlation data: it will be less when there exists a greater number of correlations which are definitely present or absent and it will be more as the number of overlapping peaks increase and the number of correlations which are uncertain increase. Similarly, the ‘quality’ of the correlation data is reflected in the number of possible solutions found: better quality data will lead to fewer solutions and poorer quality data will lead to a greater number of solutions.

Although this algorithm was developed for assigning peaks in high-resolution solid state ^{29}Si MAS NMR spectra of highly siliceous zeolites based on two-dimensional ^{29}Si INADEQUATE spectra, it may be a useful tool for easily, reliably, and efficiently working out peak assignments from other types of correlation experiments (solution or solid state) on other types of samples. In the Supplementary material, additional examples are provided in which this peak assignment algorithm is applied to solid-state dipolar-based homonuclear and heteronuclear correlation experiments performed on other inorganic solids. These examples include a ^{31}P homonuclear correlation experiment on the inorganic phosphate $\text{Cd}_3(\text{PO}_4)_2$ [14] and a combination of ^{19}F – ^{19}F , ^{31}P – ^{31}P , and ^{19}F – ^{31}P correlation

experiments on a fluorine containing gallium phosphate framework material [15].

3. Experimental

The algorithm was implemented as a *Mathematica* (version 3.0) [16] program and was run on a PC equipped with a 500 MHz Intel Pentium III processor and 256 MB of RAM. The *Mathematica* notebook and sample input files are available in the Supplementary material. The NMR spectra presented in this paper were taken from previously published work [6,7] and are used with the permission of the authors and publishers.

4. Conclusions

Using some basic graph theory concepts, a peak assignment algorithm is described which efficiently assigns the resonances in an NMR spectra to the inequivalent atoms in a structure based on the connectivity pattern between the atoms and the correlation pattern in a two-dimensional NMR experiment. An important feature of this algorithm is that it finds *all* possible sets assignments which are consistent with the experimental data in a very efficient manner. Thus when ambiguous and no simple unique assignment is possible, the availability of all the possible assignments will give a clear picture of the ambiguity and be a useful aid in planning further specific experiments to distinguish between them. Another important feature of the algorithm is that it can accommodate less than ideal experimental data in which there may be overlapping peaks and uncertainty or ambiguity about the presence or absence of correlations between the peaks.

This algorithm is particularly well-suited for situations in which there is little or no additional information to distinguish peaks from one another such as chemical shifts, relative intensities, or relaxation times. This situation often arises in solid-state extended framework systems such as siliceous zeolites, in which all of the ^{29}Si nuclei are found in very similar local tetrahedral environments, often in equal populations with similar relaxation times. This algorithm has proven to be a very useful tool for reliably assigning the peaks in the ^{29}Si MAS NMR spectra of siliceous zeolites, since peak assignment is a crucial prerequisite to determining the location of ions or organic molecules in the zeolite frameworks by double resonance solid state NMR experiments such as cross polarization and REDOR.

Supplementary material

The peak assignment algorithm (as a *Mathematica* notebook file), the input files for the examples presented

in this paper, and the additional examples described in the paper are available as Supplemental materials.

Acknowledgments

The author thanks Professor Colin Fyfe for his helpful discussions during the preparation of the manuscript and acknowledges the Natural Sciences and Engineering Research Council of Canada for financial assistance in the form of a post-graduate fellowship.

References

- [1] H.N.B. Moseley, G.T. Montelione, Automated analysis of NMR assignments and structures for proteins, *Curr. Opin. Struct. Biol.* 9 (1999) 635–642.
- [2] C. Peng, S. Yuan, C. Zheng, Y. Hui, Efficient application of 2D NMR correlation information in computer-assisted structure elucidation of complex natural products, *J. Chem. Inf. Comput. Sci.* 34 (1994) 805–813.
- [3] C.A. Fyfe, D.H. Brouwer, A.R. Lewis, J.M. Chézeau, Location of the fluoride ion in tetrapropylammonium fluoride silicalite-1 determined by $^1\text{H}/^{19}\text{F}/^{29}\text{Si}$ triple resonance CP, REDOR, and TEDOR NMR experiments, *J. Am. Chem. Soc.* 123 (2001) 6882–6891.
- [4] C.A. Fyfe, A. Diaz, A.R. Lewis, J.M. Chezeau, H. Grondey, G.T. Kokotailo, in: J.J. Fitzpatrick (Ed.), *Solid State NMR Spectroscopy of Inorganic Materials*, ACS Symposium Series, 717 (1999) 283–304.
- [5] A. Bax, R. Freeman, S.P. Kempell, Natural abundance ^{13}C – ^{13}C coupling observed via double-quantum coherence, *J. Am. Chem. Soc.* 102 (1980) 4849–4851.
- [6] C.A. Fyfe, Y. Feng, H. Gies, H. Grondey, G.T. Kokotailo, Natural-abundance two-dimensional solid state ^{29}Si NMR investigation of three-dimensional lattice connectivities in zeolite structures, *J. Am. Chem. Soc.* 112 (1990) 3264–3270.
- [7] C.A. Fyfe, H. Grondey, Y. Feng, G.T. Kokotailo, Natural-abundance two-dimensional ^{29}Si MAS NMR investigation of the three-dimensional bonding connectivities in the zeolite catalyst ZSM-5, *J. Am. Chem. Soc.* 112 (1990) 8812–8820.
- [8] C.A. Fyfe, Y. Feng, H. Grondey, G.T. Kokotailo, A. Mar, Natural abundance two-dimensional ^{29}Si MAS NMR investigation of the three-dimensional bonding connectivities in the high- and low-temperature forms of zeolite ZSM-11, *J. Phys. Chem.* 95 (1991) 3747–3751.
- [9] C.A. Fyfe, Y. Feng, H. Grondey, Evaluation of chemical shift-structure correlations from a combination of X-ray diffraction and 2D MAS NMR data for highly siliceous zeolite frameworks, *Microporous Mater.* 1 (1993) 393.
- [10] C.A. Fyfe, H. Gies, G.T. Kokotailo, B. Marler, D.E. Cox, Crystal structure of Silica-ZSM-12 by the combined use of high-resolution solid state NMR spectroscopy and synchrotron X-ray powder diffraction, *J. Phys. Chem.* 94 (1990) 3718–3721.
- [11] H. van Koningsveld, High-temperature (350 K) orthorhombic framework structure of zeolite H-ZSM-5, *Acta Cryst. B* 46 (1990) 731–735.
- [12] C.A. Fyfe, H. Grondey, Y. Feng, G.T. Kokotailo, Investigation of the three-dimensional Si–O–Si connectivities in the monoclinic form of zeolite ZSM-5 by two-dimensional ^{29}Si INADEQUATE experiments, *Chem. Phys. Lett.* 173 (1990) 211–215.

- [13] H. van Koningsveld, J.C. Jansen, H. van Bekkum, The monoclinic framework structure of zeolite H-ZSM-5. Comparison with the orthorhombic framework of as-synthesized ZSM-5, *Zeolites* 10 (1990) 235–242.
- [14] W.A. Dollase, M. Feike, H. Förster, T. Scaller, I. Schnell, A. Sebal, S. Steuernagel, A 2D ^{31}P MAS NMR study of polycrystalline $\text{Cd}_3(\text{PO}_4)_2$, *J. Am. Chem. Soc.* 119 (1997) 3807–3810.
- [15] V. Munch, F. Taulelle, T. Loiseau, G. Férey, A.K. Cheetham, S. Weigel, G.D. Stucky, A set of two-dimensional solid-state NMR experiments providing topological information for structure refinement of TMP-GaPO, *Magn. Reson. Chem.* 37 (1999) S100–S107.
- [16] S. Wolfram, *Mathematica: A System for Doing Mathematics by Computer* v. 3.0, Champaign IL, Wolfram Media, 1996.