

# Chapter 5

## Multidimensional NMR by Projection-Reconstruction

Ray Freeman<sup>1</sup> and Ēriks Kupče<sup>2</sup>

<sup>1</sup>*Jesus College, Cambridge University, Cambridge, CB5 8BP, UK*

<sup>2</sup>*Varian Ltd, 6 Mead Road, Yarnton, Oxford, OX5 1QU, UK*

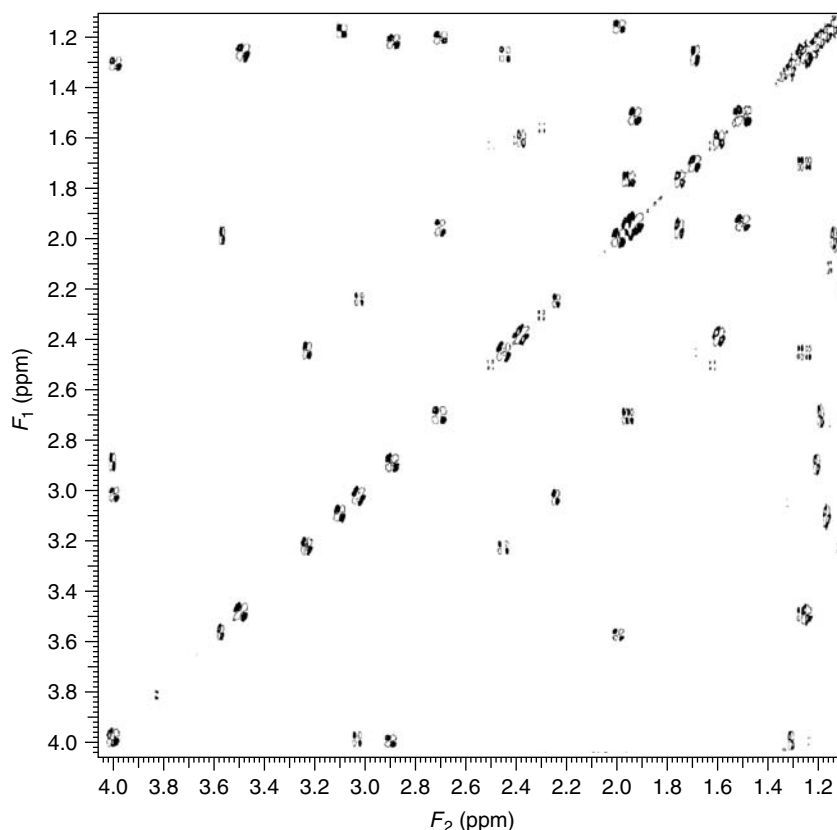
5.1	Introduction	73
5.2	Speeding Up Multidimensional Spectroscopy	74
5.3	Projection-Reconstruction	75
5.4	A Practical Illustration	80
5.5	Higher Dimensions	80
5.6	Minimal Sampling	81
5.7	Hyperdimensional Spectroscopy	82
	References	83

### 5.1 INTRODUCTION

Spectroscopy in several frequency dimensions has revolutionized the field of nuclear magnetic resonance (NMR), triggered by the initial experiments of Jeener and the extension to a broad range of applications by Ernst<sup>1</sup> (see Chapter 2). Instead of the familiar graph of absorption intensity plotted against NMR frequency, a two-dimensional spectrum displays a contour map of absorption peaks as a function of two frequency coordinates. As an example, Figure 5.1 shows the 800 MHz two-dimensional proton spectrum of a sample of erythromycin A.

This was recorded with the correlation spectroscopy (COSY) sequence (see Chapter 13). The dominant responses along the principal diagonal ( $F_1 = F_2$ ) are of little intrinsic interest, but each off-diagonal peak indicates an interaction between the two chemical sites defined by their  $F_1$  and  $F_2$  coordinates. Spectra such as this tell us which chemical sites are related by spin–spin coupling, cross-relaxation, or chemical exchange. They provide the chemist with an easily recognized picture of the interactions within the molecule under investigation, information that previously had to be obtained through a rather tedious set of double resonance experiments.

Higher-dimensional experiments are also possible, the results being displayed as plane projections of the higher-dimensional data matrix. A multidimensional experiment is set up in the time domain, with nuclear spins precessing under different conditions during a sequence of evolution periods ( $t_1, t_2, \dots$ ) before conversion to a detectable signal in a final acquisition interval  $t_A$ . Fourier transformation with respect to  $t_1, t_2, \dots, t_A$  generates a multidimensional spectrum in the corresponding frequency space,  $F_1, F_2, \dots, F_A$ . Initially applied to small organic molecules such as erythromycin, multidimensional NMR is now used for structural and dynamic studies of proteins and other large biomolecules. The additional dimensions are used to monitor the evolution of heteronuclear spins, notably  $^{13}\text{C}$  or  $^{15}\text{N}$ , or to resolve the overlap



**Figure 5.1.** The 800 MHz two-dimensional correlation double-quantum filtered COSY (see Chapter 17) spectrum of erythromycin A. Peaks on the principal diagonal represent proton signals that do not change frequency during the experiment. Off-diagonal peaks indicate transfer of proton magnetization between the chemical sites defined by the frequency coordinates  $F_1$  and  $F_2$ . (Spectrum courtesy of Dr. Toshiaki Nishida.)

problems that often occur in the spectra of large molecules. Today no self-respecting chemical or biochemical laboratory can afford to operate without multidimensional NMR.

## 5.2 SPEEDING UP MULTIDIMENSIONAL SPECTROSCOPY

As more and more complex spectra are investigated, the trend is to introduce higher and higher dimensions, particularly in the booming discipline of biological NMR spectroscopy. It soon became apparent that the duration of high-dimensional experiments was a severe limitation, as the standard protocol requires each and every evolution dimension

to be comprehensively sampled in order to satisfy the Nyquist condition and the resolution requirements. For example, a four-dimensional experiment requiring each of the three evolution dimensions to be explored independently with 64 time increments followed by a 1 s period for signal acquisition and relaxation, would have an overall duration of three days. Even with severe compromises on the sampling requirements, protein studies often tie up an expensive spectrometer for many days.

Several palliative solutions have been explored. The simplest of these involves fast pulsing with Ernst-angle excitation, or polarization sharing with adjacent nuclei. A similar modest improvement in efficiency stems from the recent introduction of four independent receivers tuned to (say)  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ , and  $^{31}\text{P}$  at the same time. However, most of

the innovations rely on the concept of *sparse sampling* of evolution space instead of the systematic step-wise exploration of every coordinate on a Cartesian grid. For example, the evolution interval may be prematurely truncated and then artificially extended by a linear prediction algorithm (see Chapter 10), saving a factor of about two in each evolution dimension. Regular sampling can be replaced by exponentially-weighted sampling where the outer regions of evolution space are relatively poorly defined, or indeed by random sampling. It is possible to discover (and exploit) the minimum number of samples needed to obtain a multidimensional spectrum without aliasing or compromising the spectral resolution. Because of its unorthodox approach, the filter diagonalization method (see Chapter 9) attacks the problem by using the fine digitization in the acquisition dimension to compensate for restricted sampling in the evolution domain. Each of these limited sampling schemes engenders some kind of spectral artifacts. These are more readily accommodated if they are well defined. Sparse sampling by projection-reconstruction is one such method, and this is the theme of the next section.

It is important to remember that the extensive sampling employed in a conventional multidimensional experiment improves the signal-to-noise ratio in proportion to the square root of the number of samples. Sparse sampling methods inevitably sacrifice part of this advantage, and are therefore only applicable to situations where the intrinsic sensitivity is already acceptable. This is called the *sampling-limited* case, as opposed to the *sensitivity-limited* situation. Fortunately, the signal-to-noise ratio achieved in modern NMR spectrometers, particularly those equipped with a cryogenically-cooled receiver coil, does not normally require protracted time averaging.

### 5.3 PROJECTION-RECONSTRUCTION

Auguste Rodin (1840–1917) was probably the most famous and controversial sculptor at the turn of the twentieth century. His sculptures were so lifelike that his critics accused him of cheating by taking direct plaster casts of his models. In fact, he had developed a rather original method for converting what he observed into the finished product. The model sat on a rotating turntable with backlighting so that Rodin saw only a two-dimensional silhouette. He then rotated the turntable in small angular steps to

obtain a set of different views, thereby gathering the information needed to construct a faithful three-dimensional image. A case can be made that Rodin was the true originator of projection reconstruction.

Now if we think of a three-dimensional NMR spectrum as an *object*, it follows that it should be possible to reconstruct it from a suitable set of plane projections, just as a three-dimensional image of the human body is reconstructed from projections of the X-ray absorption in the X-ray scanner, or from NMR absorption in magnetic resonance imaging. Indeed, the spectroscopic case should be more tractable, as an NMR spectrum is a relatively sparse distribution of discrete resonances, rather like the stars in the night sky, compared with the continuous absorption characteristic of the scanner. The standard method for obtaining a three-dimensional NMR spectrum is to record a new scan for each and every increment of the evolution variables,  $t_1$  and  $t_2$ , satisfying the Nyquist sampling condition and the resolution requirements in each evolution dimension. As mentioned above, it can involve a protracted measurement, particularly for the higher-dimensional versions. By contrast, several plane projections at different inclinations can be recorded relatively quickly, hence the speed advantage of projection reconstruction. The more sparsely we sample evolution space, the sooner the measurement is complete. This may permit investigation of unstable biological molecules, or chemical exchange, or even, in very favorable cases, protein folding. However, the main goal is to bring down the experimental duration to practically acceptable times. The experimental duration is set, not by signal-to-noise considerations, but by the sampling requirements. Sensitivity is traded for speed. An  $N$ -dimensional projection reconstruction experiment enjoys a speed advantage of roughly  $N-2$  orders of magnitude compared with that of the conventional protocol.

There is, of course, a price to pay for this speed advantage. Although the sensitivity per unit time remains essentially unchanged, the overall sensitivity is reduced in proportion to the time factor, as in all NMR experiments. When sensitivity is marginal, some of the speed advantage can be sacrificed, yet still leave the projection reconstruction scheme with an appreciable speed advantage. In the limit of very poor intrinsic sensitivity, sparse sampling offers no practical advantage, and the traditional Cartesian sampling scheme is the most efficient option.

### 5.3.1 Projection

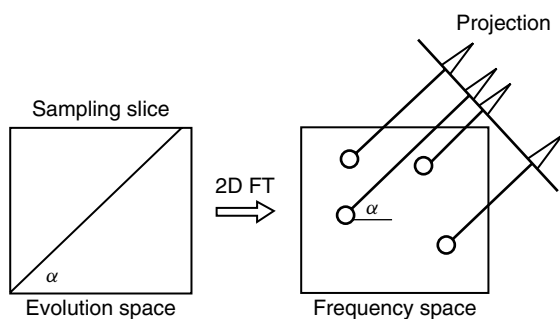
Recording a set of plane projections of a three-dimensional NMR spectrum is a relatively straightforward process, relying on an extension of Bracewell's "slice/projection" Fourier transform theorem.<sup>2</sup> For clarity, it is convenient to focus attention on the two dimensions ( $t_1$  and  $t_2$ ) of evolution space (Figure 5.2). Fourier transformation of the signal recorded along a slice through this space, passing through the origin at an inclination  $\alpha$  with respect to the  $t_1$  axis, yields a frequency-domain signal  $S(F_1, F_2)$  projected at the same angle  $\alpha$  with respect to the  $F_1$  axis. If the evolution times are incremented in steps  $\Delta t_1$  and  $\Delta t_2$ , then  $\alpha = \arctan(\Delta t_2/\Delta t_1)$ . The relative scaling factor determines the angle of the tilted projection. It was the "accordion" experiment<sup>3</sup> that first exploited the idea of linking the incrementation of  $t_1$  and  $t_2$  to derive skew projections.

If coherence evolves during  $t_1$  at a frequency  $\Omega_A$  and then during  $t_2$  at a frequency  $\Omega_B$ , then the signal detected during the acquisition interval  $t_3$  is doubly modulated. The two evolution parameters are linked according to

$$t_1 = t \cos \alpha \quad (5.1)$$

$$t_2 = t \sin \alpha \quad (5.2)$$

By shifting the phases of the coherence transfer pulses ( $0^\circ 0^\circ$ ,  $0^\circ 90^\circ$ ,  $90^\circ 0^\circ$ ,  $90^\circ 90^\circ$ ) four



**Figure 5.2.** The Bracewell slice/projection theorem. Fourier transformation of a slice through the two-dimensional time-domain signal at an arbitrary angle  $\alpha$  generates a projection of the corresponding frequency-domain spectrum at the same angle  $\alpha$ . In this manner, a set of differently oriented projections can be recorded and used to reconstruct the full spectrum.

combinations of real and imaginary components are recorded:

$$S_1 = \cos(\Omega_A t \cos \alpha) \cos(\Omega_B t \sin \alpha) \quad (5.3)$$

$$S_2 = \cos(\Omega_A t \cos \alpha) \sin(\Omega_B t \sin \alpha) \quad (5.4)$$

$$S_3 = \sin(\Omega_A t \cos \alpha) \cos(\Omega_B t \sin \alpha) \quad (5.5)$$

$$S_4 = \sin(\Omega_A t \cos \alpha) \sin(\Omega_B t \sin \alpha) \quad (5.6)$$

Standard trigonometric transformations yield the expressions:

$$S_1 - S_4 = \cos(\Omega_A \cos \alpha + \Omega_B \sin \alpha)t \quad (5.7)$$

$$S_2 + S_3 = \sin(\Omega_A \cos \alpha + \Omega_B \sin \alpha)t \quad (5.8)$$

$$S_1 + S_4 = \cos(\Omega_A \cos \alpha - \Omega_B \sin \alpha)t \quad (5.9)$$

$$S_2 - S_3 = \sin(\Omega_A \cos \alpha - \Omega_B \sin \alpha)t \quad (5.10)$$

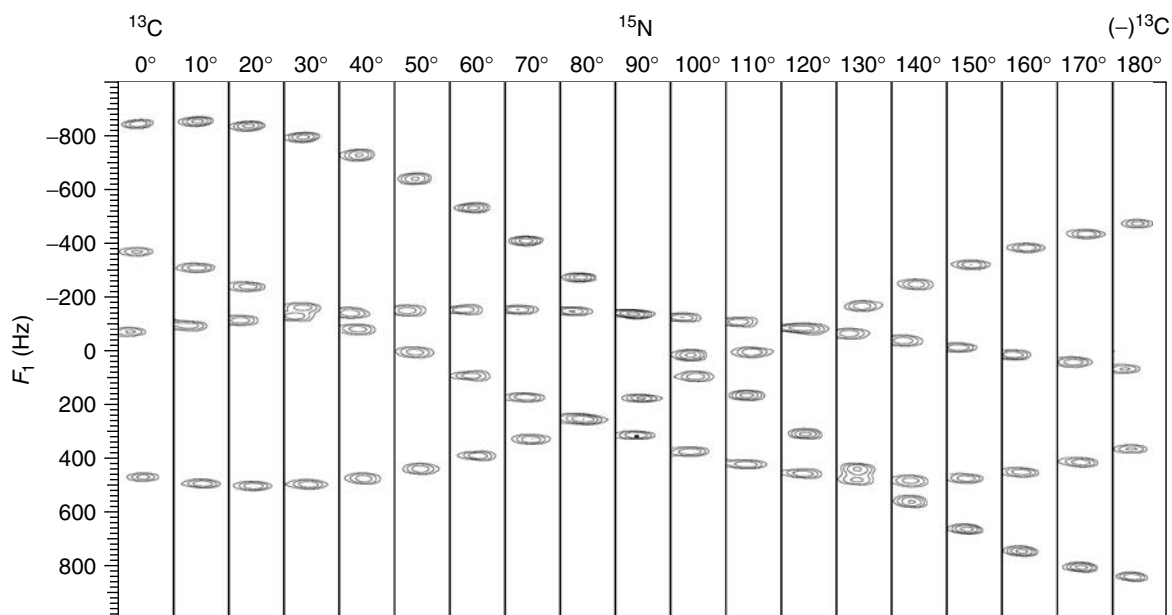
As a consequence, hypercomplex Fourier transformation generates two frequencies:

$$\Omega(+\alpha) = \Omega_A \cos \alpha + \Omega_B \sin \alpha \quad (5.11)$$

$$\Omega(-\alpha) = \Omega_A \cos \alpha - \Omega_B \sin \alpha \quad (5.12)$$

Consequently, projections are recorded in pairs at  $\pm\alpha$ , providing two independent views of the three-dimensional spectrum. The  $\cos \alpha$  and  $\sin \alpha$  terms reflect the extent to which the projection plane is tilted with respect to the  $F_1 F_3$  plane. Figure 5.3 shows the sinusoidal dependence of the experimental frequencies recorded for the three-dimensional HNCQ spectrum of ubiquitin projected onto planes progressively tilted in  $10^\circ$  increments, starting with the carbon–proton plane ( $\alpha = 0^\circ$ ) running through the nitrogen–proton plane ( $\alpha = 90^\circ$ ) and on toward the reversed carbon–proton plane ( $\alpha = 180^\circ$ ).<sup>4</sup>

The special cases  $t_1 = 0$  ( $\alpha = 90^\circ$ ) and  $t_2 = 0$  ( $\alpha = 0^\circ$ ) are well known to give the "first planes"  $S(F_2 F_3)$  and  $S(F_1 F_3)$ , respectively, and are commonly used in the setting-up protocol for three-dimensional measurements. Ding and Gronenborn,<sup>5</sup> Kim and Szyperski,<sup>6</sup> and Kozsminski and Zhukov<sup>7</sup> have exploited another special case where  $t_1$  and  $t_2$  are incremented at identical rates, deriving sum and difference frequencies that can be measured directly, because then the  $\cos \alpha$  and  $\sin \alpha$  terms are equal. For projection-reconstruction NMR, this is an unnecessary restriction; there is no particular reason to choose  $\alpha = 45^\circ$ . A more general scheme is used, employing a set of differently inclined projections, usually including the two first planes  $S(F_2 F_3)$  and  $S(F_1 F_3)$ .



**Figure 5.3.** The sinusoidal dependence of the frequencies observed in the three-dimensional HNCOS spectrum of ubiquitin correlating the  $^{15}\text{N}$ ,  $^{13}\text{C}$ , and  $^1\text{H}$  resonances. By changing the relative rates of incrementation of the evolution time  $t_1$  and  $t_2$ , we obtain a frequency-domain spectrum in a plane progressively tilted in  $10^\circ$  steps, starting at the  $^{13}\text{C}$ – $^1\text{H}$  plane, moving through the  $^{15}\text{N}$ – $^1\text{H}$  plane, and on to the reversed  $^{13}\text{C}$ – $^1\text{H}$  plane. This implies that frequency degeneracies that may occur in a two-dimensional spectrum can be readily avoided in a tilted three-dimensional spectrum. (Reproduced from Ref. 4. © The American Chemical Society, 2004.)

In an ideal world, the slices through evolution space would be taken at equal angular steps, but in crowded three-dimensional spectra, occultation may be a complication if two or more resonances lie along the same projection ray. The general solution is to employ more projections, reducing the incidence of accidental alignment. The penalty is that this reduces the speed advantage compared with full systematic sampling on a Cartesian grid. However, there is a way to predict which projection angles avoid occultation and thus are more suitable. Consider the first planes,  $S(F_2F_3)$  and  $S(F_1F_3)$ , normally the first measurements to be performed. A two-dimensional convolution of these spectra creates a *provisional* three-dimensional lattice of cross-peaks, some of which are genuine and some false. This permits a quick calculation of the risk of occultation as a function of the projection angle (Figure 5.4). It overestimates the danger (by involving false cross-peaks) but, nevertheless, it indicates all those inclinations that avoid accidental alignment, suggesting the best

choice of  $\alpha$  for the next projection measurement. An adaptive program can then be written to set up the minimum number of projections needed to obtain an unambiguous reconstruction.

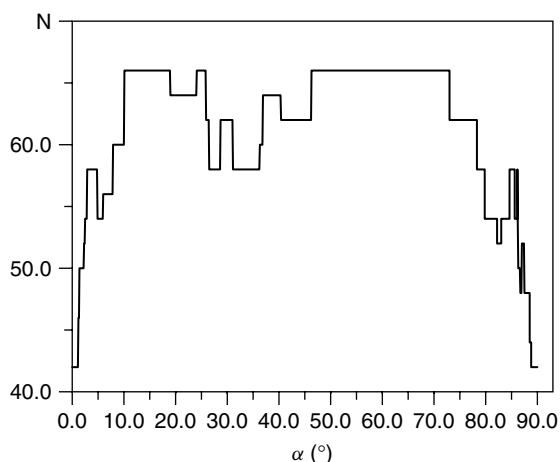
### 5.3.1.1 Linewidths

In a three-dimensional experiment, the various evolving nuclear spins may have significantly different time constants for spin–spin relaxation. Suppose that these are  $T_2^A$  in the  $t_1$  interval and  $T_2^B$  in the  $t_2$  interval. The decay of magnetization during evolution can then be expressed as

$$M(t) = M(0) \exp \left[ -t \left( \frac{\cos \alpha}{T_2^A} + \frac{\sin \alpha}{T_2^B} \right) \right] \quad (5.13)$$

And an effective spin–spin relaxation time can be defined as

$$T_2^{\text{eff}} = \frac{T_2^A T_2^B}{(T_2^B \cos \alpha + T_2^A \sin \alpha)} \quad (5.14)$$



**Figure 5.4.** Selecting a suitable projection angle  $\alpha$ . The  $0^\circ$  and  $90^\circ$  projections of the 700 MHz spectrum of agitoxin were used to reconstruct a provisional three-dimensional spectrum containing both genuine and false correlation peaks. This allows us to calculate when two or more peaks are eclipsed in the projection (indicated by dips in the graph). There are appreciable ranges of  $\alpha$  that avoid this problem. (Reproduced from *Magn. Reson. Chem.* **44**, 197. © John Wiley & Sons Ltd, 2006.)

This corresponds to an effective Lorentzian line with full width at half-height:

$$\Delta\nu = \frac{1}{(\pi T_2^{\text{eff}})} = \frac{\cos \alpha}{(\pi T_2^A)} + \frac{\sin \alpha}{(\pi T_2^B)} \quad (5.15)$$

This indicates that in a tilted projection, the linewidths are broader than those in at least one of the first planes  $S(F_2F_3)$  and  $S(F_1F_3)$ , suggesting that if  $T_2^A$  is longer than  $T_2^B$ , it may be worthwhile selecting a small tilt angle  $\alpha$  to obtain narrower lines in this projection.

### 5.3.2 Reconstruction

Reconstruction methods are employed in many fields—geophysics, ultrasonics, X-ray tomography, and magnetic resonance imaging. On a more mundane level, information about an everyday object can be converted into a set of projections that are transmitted digitally and then reconverted into a three-dimensional image that can be viewed from any angle; the act of advertising a used car for sale might well make use of this internet technology.

Although image reconstruction is widely used, it is not entirely straightforward, being formally represented by the *inverse* Radon transform. The image density in the plane under consideration is given by

$$f(x, y) = \int p(x \cos \alpha + y \sin \alpha, \alpha) d\alpha \quad (5.16)$$

where  $p$  is the perpendicular distance of the ray from the origin and  $\alpha$  is the angle of incidence of the ray. The task is to find a three-dimensional density function compatible with all the experimental projections. The approaches may be deterministic (basically back-projection) or statistical (iterative fitting of a model). First consider the deterministic schemes. The choice between three possible back-projection methods depends on the intrinsic sensitivity and the complexity of the NMR spectrum.

#### 5.3.2.1 Spectra with Poor Sensitivity: The Additive Algorithm

The problem is most easily visualized in terms of a two-dimensional reconstruction; the three-dimensional version would simply be an assembly of a set of reconstructed planar images stacked parallel to one another. Suppose that there are  $P$  projections in all. Each individual projection trace is back-projected at right angles across the image plane as a set of parallel rays. The intensity at any given pixel in this plane is given by the sum of the intensities of  $P$  intersecting rays, one from each projection. The merit of this “additive algorithm” is that the signal-to-noise ratio increases as  $\sqrt{P}$ . Furthermore, if sensitivity is so low that a particular resonance is missing from a particular projection, the reconstruction is not compromised. Note that it is better to enhance the signal-to-noise ratio by increasing  $P$  rather than to employ multiscan averaging of several projections at the same angle of incidence.

The disadvantage of this additive algorithm is that it generates artifacts in the form of spurious ridges that criss-cross the image plane. An intersection of these ridges defines a correlation peak, but this may be either genuine or false. A genuine peak is identified as the point where ridges from all  $P$  projections intersect, but where fewer than  $P$  ridges cross, a false peak appears. (Of course, in the unfortunate case that a resonance line is missing from one projection, a genuine correlation peak involves only  $P-1$



intersections.) As  $P$  is increased, the intensities of true image peaks are enhanced with respect to the false peaks, and the undesirable ridges, although now more numerous, become relatively weaker and tend eventually to sink below the level of the baseplane noise. The quality of the reconstructed spectrum is improved, not only because of multiscan averaging, but also through artifact reduction. This algorithm, therefore, favors the use of a relatively large number of independent projections. Additive back-projection of resonances with appreciable linewidths leads to an undesirable broadening of the skirts of the reconstructed correlation peak, but this can be remedied by “filtered back-projection”, which de-emphasizes the wings of the resonance lines in the projections.

If only a few projections are used, leaving some obtrusive artifacts, a “CLEAN” algorithm adapted from radioastronomy can be applied to remedy the problem. It works by finding the tallest correlation peak and then subtracting it from the reconstructed image, along with its associated ridges. An iterative program then takes out the remaining peaks and their pattern of ridges one at a time. The search is stopped just above the level of the baseplane noise. The correlation peaks are then reintroduced into the image stripped of their associated ridges, giving an artifact-free reconstruction. Once the ridges have been eliminated, there are no false correlation peaks. The noisy baseplane is retained as a direct indication of the reliability of the measurements, as artificially noise-free images can be misleading.

### 5.3.2.2 *Spectra with Good Sensitivity: The Lowest-Value Algorithm*

A very effective suppression of artifacts can be achieved by replacing additive back-projection with a “lowest-value” algorithm. The intensity of a given pixel in the image plane is determined not by the sum, but by the lowest intensity of all the intersecting rays. Where a false correlation peak would otherwise occur, one of the  $P$  contributing back-projections carries only noise; hence, the false peak is suppressed. Spurious ridges are similarly eliminated. A genuine correlation peak appears only if there is a contributing NMR response in all  $P$  projections. Good sensitivity must be assumed. If any projection contains a resonance line that is so weak as to escape detection, the corresponding correlation peak is missed, unless the discrepancy is recognized beforehand so that the defective projection can be excluded from

the reconstruction. The lowest-value algorithm is best suited to reconstructions that employ a rather small number  $P$  of projections; the signal-to-noise ratio of the reconstructed image does not increase with  $P$ . It also has the effect of slightly reducing baseplane noise, although this serves no practical advantage. A rather trivial consequence of the nonlinearity of this algorithm is a slight distortion of the reconstructed lineshapes, giving the intensity contours some polygonal character (with  $2P$  edges).

If a relatively large number of projections are recorded, it is possible to use a hybrid scheme that exploits *both* the benefits of the lowest-value algorithm (lack of artifacts) *and* those of the additive algorithm (sensitivity) to get the best of both worlds. For example, a certain degree of artifact suppression is achieved by processing small independent subgroups of projections with the lowest-value algorithm to derive several different versions of the reconstructed image. Then coaddition of these images tends to emphasize genuine correlation peaks at the expense of false peaks.

### 5.3.2.3 *Very Crowded Spectra: The Algebraic Algorithm*

Multidimensional NMR studies of biological macromolecules such as proteins often have to deal with complex crowded spectra. It is then unrealistic to retain linewidth and lineshape information for every response; indeed, it is often necessary to forego intensity information as well, concentrating only on frequencies. This implies the application of a peak-picking routine to convert each experimental projected trace into a frequency list. The danger lies concealed in this initial processing stage. If two NMR responses in the projected trace are unresolved, peak-picking finds only one frequency, even when it is clear on inspection that there is a second line on the shoulder of the first. Similarly, a very weak response may be lost forever if it lies below the arbitrary intensity threshold set by the peak-picking algorithm. These uncertainties complicate the reconstruction stage.

Once each projection has been reduced to a frequency list, reconstruction becomes an exercise in algebra. A typical frequency defines an equation representing a straight line running across the image plane at the appropriate inclination  $\alpha$ . In an ideal world, the criterion for correlation would be that  $P$  such lines, *one from each projection*, intersect within

a specified margin of error. Intersections involving fewer than  $P$  lines would be dismissed as false correlations. In actual practice, when some frequencies may not have been identified by the peak-picking routine, a genuine correlation can involve fewer than  $P$  simultaneous equations. This seriously complicates the problem of distinguishing between true and false. The solution is to exploit a version of the CLEAN algorithm. As each  $P$ -fold intersection is identified, the contributing frequencies are removed from the frequency lists in an iterative program that continues until all such  $P$ -fold correlations have been discovered. This has the effect of reducing the probability of finding false correlations. In the next stage, solutions involving only  $P-1$  intersections are identified and removed one by one, progressively lowering the incidence of false solutions. Further iterative stages may be needed, investigating cases of lower and lower-order intersections. The advantage of this algebraic algorithm is the vast reduction in the volume of data to be processed, a point of particular importance when dealing with biomolecules such as proteins.

#### 5.3.2.4 Statistical Methods of Reconstruction

The *inverse* nature of the reconstruction stage is highlighted in schemes that employ statistical methods to fit a provisional model of the three-dimensional spectrum to the experimental projection data. These methods include iterative programs widely used in other fields—least-squares fitting, maximum entropy, maximum likelihood, or Bayesian inference. The speed of these computations is significantly increased if some prior knowledge can be provided, perhaps from an initial deterministic analysis or by making assumptions about the expected NMR lineshapes. As model-fitting methods are fundamentally nonlinear operations, the final result may be a clean, noise-free reconstruction, but this apparent advantage is illusory. Random noise present in the experimental projections cannot be realistically explored by the fitting program; hence, weak NMR responses of the same order as the baseline noise are necessarily overlooked. Worse still, experimental artifacts may be promoted to the status of true NMR responses because of an initial assumption that all the image features have (say) Gaussian lineshapes. Because of the possibility of such false negatives and false positives, statistical reconstruction methods should be used with particular caution.

## 5.4 A PRACTICAL ILLUSTRATION

Projection-reconstruction has been used with some success for the study of protein samples that traditionally have required long-duration NMR measurements. An illustrative example is the 800 MHz three-dimensional spectrum of  $^{15}\text{N}$ - and  $^{13}\text{C}$ -enriched nuclease A inhibitor reconstructed from projections recorded at  $0^\circ$ ,  $\pm 30^\circ$ ,  $\pm 60^\circ$ , and  $90^\circ$ . The pulse sequence (HNCA) is widely used to provide aminoacid sequencing by correlating the amide protons and  $^{15}\text{N}$  sites with the intraresidue  $\text{C}\alpha$  site via the direct  $^{15}\text{N}$ – $^{13}\text{C}$  coupling, and with the  $\text{C}\alpha$  site of the preceding residue via the two-bond  $^{15}\text{N}$ – $^{13}\text{C}$  coupling. We can then display the results as a “walk-through” of adjacent residues<sup>4</sup> in the form of connected strip plots, one for each proton site. Figure 5.5 maps the nine  $^{15}\text{N}$ – $^{13}\text{C}$  correlations along a chain F37 to A46 terminated by proline residues (which have no NH groups). The experiment required 10 min per projection corresponding to a 1 h measurement, whereas the estimated duration of a conventional experiment would be 11 h.

## 5.5 HIGHER DIMENSIONS

The same principles can be used to reconstruct multi-dimensional spectra in more than three dimensions.<sup>4</sup> This strategy becomes increasingly important for the crowded spectra of biological molecules, where accidental overlap is common in three-dimensional spectra, and information about  $^{13}\text{C}$  and  $^{15}\text{N}$  spins is often employed. For example, a four-dimensional measurement, involving three successive evolution intervals  $t_1$ ,  $t_2$ , and  $t_3$ , allows them to be linked in pairs or all together. A second tilt angle  $\beta$  is introduced, and the linked evolution times are now defined as

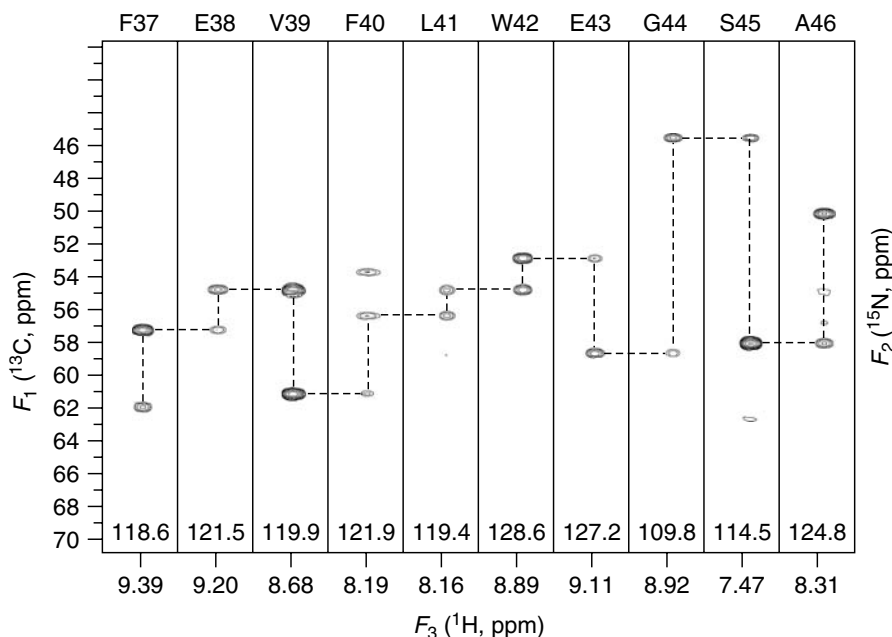
$$t_1 = t \cos \alpha \cos \beta \quad (5.17)$$

$$t_2 = t \sin \alpha \cos \beta \quad (5.18)$$

$$t_3 = t \sin \beta \quad (5.19)$$

The problem is best visualized in three-dimensional “evolution space”  $F_1 F_2 F_3$  where  $F_4$  is the acquisition dimension. The tilt angle  $\alpha$  represents a rotation about the  $F_3$  axis, whereas the new tilt angle  $\beta$  represents a general rotation about an axis in the  $F_1 F_2$  plane. Projections are onto one of the orthogonal planes,  $F_1 F_4$  (when  $\alpha = \beta = 0^\circ$ ),  $F_2 F_4$  (when





**Figure 5.5.** Strip plots from the 800 MHz three-dimensional HNCA spectrum of nuclease A inhibitor (3 mM aqueous solution), showing the chain of nine residues correlated through their  $^{15}\text{N}$  and  $^{13}\text{C}$  resonances. The chain from F37 to A46 is terminated by proline residues (which have no NH groups). The lowest-value reconstruction algorithm was used. The total duration of the measurement was 1 h compared with an estimated 11 h for the conventional mode, using 64 and 32 time-domain increments. (Reproduced from Ref. 4. © The American Chemical Society, 2004.)

$\alpha = 90^\circ$ ,  $\beta = 0^\circ$ ), or  $F_3F_4$  (when  $\alpha = 0^\circ$ ,  $\beta = 90^\circ$ ) or onto singly tilted planes (when either  $\alpha$  or  $\beta$  is zero) or onto doubly tilted planes (when neither  $\alpha$  nor  $\beta$  is zero). Doubly tilted projections correspond to the case that  $t_1$ ,  $t_2$ , and  $t_3$  are all linked together. If the evolving frequencies are  $\Omega_A$ ,  $\Omega_B$ , and  $\Omega_C$ , respectively, then after hypercomplex Fourier transformation these doubly tilted projections yield four combinations of frequencies:

$$\Omega_A \cos \alpha \cos \beta + \Omega_B \sin \alpha \cos \beta + \Omega_C \sin \beta \quad (5.20)$$

$$\Omega_A \cos \alpha \cos \beta + \Omega_B \sin \alpha \cos \beta - \Omega_C \sin \beta \quad (5.21)$$

$$\Omega_A \cos \alpha \cos \beta - \Omega_B \sin \alpha \cos \beta + \Omega_C \sin \beta \quad (5.22)$$

$$\Omega_A \cos \alpha \cos \beta - \Omega_B \sin \alpha \cos \beta - \Omega_C \sin \beta \quad (5.23)$$

Reconstruction may involve tilted and doubly tilted projections, but the recorded spectra are plane projections of the full four-dimensional data matrix. The extension to five dimensions follows the same principles, but the formal treatment is correspondingly more complex.

## 5.6 MINIMAL SAMPLING

The on-going quest for fast results raises the question—is there a minimum sampling regime that still satisfies the resolution requirements? Consider the three-dimensional case. The first planes  $F_1F_3$  and  $F_2F_3$  are obtained by setting  $t_2 = 0$  and  $t_1 = 0$ , respectively. These two-dimensional measurements have relatively high sensitivity because relaxation and coherence transfer losses are limited to just one evolution period. Since these planes provide accurate information about all the evolving frequencies, they can, in principle, be used to derive the full three-dimensional matrix. In practice, such a reconstruction would be compromised if there were degeneracies in the chemical shifts. Exact overlap of shifts in the  $F_1$  or  $F_2$  dimension remains unresolved, even with full conventional three-dimensional methodology, unless further frequency dimensions are introduced. This leaves one remaining source of ambiguity—overlapping chemical shifts in the common frequency dimension  $F_3$ . Consider the

example of  $^{13}\text{C}$ – $^{15}\text{N}$ – $^1\text{H}$  correlation spectroscopy, where  $^{13}\text{C}$  spins precess at  $\Omega_{\text{C}}$  during  $t_1$  and  $^{15}\text{N}$  spins at  $\Omega_{\text{N}}$  during  $t_2$ , with proton acquisition. In the simple case where all the proton responses are resolved, the three-dimensional spectrum can be reconstructed without ambiguity, as each  $^{13}\text{C}$  site is correlated with the appropriate  $^{15}\text{N}$  site through coupling to the same proton. However, if two proton responses overlap, then there are two evolving  $^{13}\text{C}$  frequencies  $\Omega_{\text{C}}(1)$  and  $\Omega_{\text{C}}(2)$  and two  $^{15}\text{N}$  frequencies  $\Omega_{\text{N}}(1)$  and  $\Omega_{\text{N}}(2)$ . The ambiguity arises because  $\Omega_{\text{C}}(1)$  may correlate either with  $\Omega_{\text{N}}(1)$  or with  $\Omega_{\text{N}}(2)$ . A single new measurement is now carried out at fixed values ( $t_1^*$  and  $t_2^*$ ) of the evolution times.<sup>8</sup> This is essentially a one-dimensional experiment and is very fast. There are two possible expressions for the detected proton intensity:

$$P \left\{ \cos [\Omega_{\text{C}}(1)t_1^*] \cos [\Omega_{\text{N}}(1)t_2^*] + [\Omega_{\text{C}}(2)t_1^*] \right. \\ \left. \times \cos [\Omega_{\text{N}}(2)t_2^*] \right\} \quad (5.24)$$

$$P \left\{ \cos [\Omega_{\text{C}}(1)t_1^*] \cos [\Omega_{\text{N}}(2)t_2^*] + [\Omega_{\text{C}}(2)t_1^*] \right. \\ \left. \times \cos [\Omega_{\text{N}}(1)t_2^*] \right\} \quad (5.25)$$

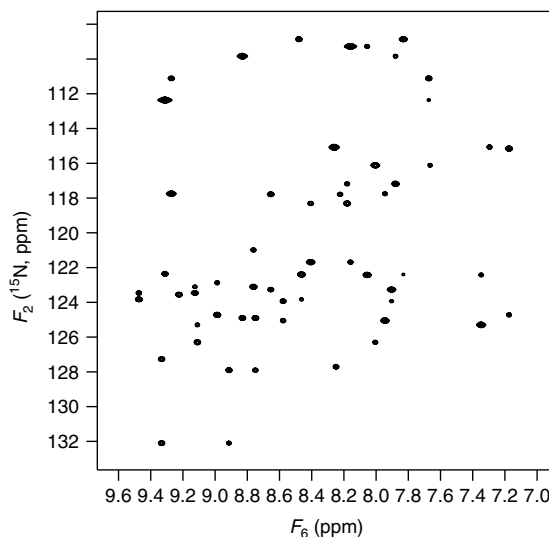
These can be discriminated because accurate values for the frequencies  $\Omega_{\text{C}}(1)$ ,  $\Omega_{\text{C}}(2)$ ,  $\Omega_{\text{N}}(1)$ , and  $\Omega_{\text{N}}(2)$  are known from the first planes. These *calculated* intensity values are now compared with the *experimental* values to find the best match. This has been called “3D-SPEED” (single-point evaluation of the evolution domain) and can be extended to higher-dimensional experiments (see below).

## 5.7 HYPERDIMENSIONAL SPECTROSCOPY

If we focus attention on two adjacent aminoacid residues of a protein, the backbone  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  nuclei form a coupled 10-spin system, and in an ideal world we might be tempted to run a 10-dimensional experiment in order to pin down all 45 pairwise interactions. In fact, this is quite impractical, because even with only 16 complex data points in each of the nine evolution intervals and one second for signal acquisition and spin-lattice relaxation, the expected duration would be  $2^{45}$  s (a million years). Furthermore, there would be unacceptable signal losses due to the cumulative effects of pulse imperfections, spin–spin

relaxation, and a  $\sqrt{2}$  attenuation at each coherence transfer stage.

A solution is to postulate a *virtual* 10-dimensional data matrix, never reconstructed in its entirety, but used as an imaginary scaffold on which to erect low-dimensional spectra obtained by standard pulse sequences. It is perfectly feasible to combine the appropriate low-dimensional spectra to generate a higher-dimensional spectrum. For example, a three-dimensional spectrum can be reconstructed from two two-dimensional spectra that share a common frequency axis, provided that precautions are taken to resolve any frequency degeneracy on the common axis (see above). In this manner we gradually build up a hyperdimensional spectrum that may then be projected onto the appropriate planes to give the desired correlations. We reconstruct only those parts of the virtual hyperdimensional data matrix that are required by the structural problem under investigation.



**Figure 5.6.** Hyperdimensional NMR. A five-dimensional subspectrum has been built up from a combination of three-dimensional HNCA and HN(CO)CA spectra. This forms part of the *virtual* 10-dimensional 800 MHz spectrum of agitoxin. The projection onto the  $F_6$ – $F_2$  plane shows the correlations between  $\text{N}(i-1)$  and  $\text{NH}(i)$  sites. This implies that any of the 45 possible pairwise correlations can be derived by analogous combinations of low-dimensional measurements. (Reproduced from Ref. 9. © Elsevier, 2008.)

A practical example<sup>9</sup> is provided by the investigation of adjacent aminoacid residues in a small 39-residue protein, agitoxin, isotopically enriched in <sup>15</sup>N and <sup>13</sup>C. The spectra were recorded at 800 MHz on a 0.3 mM aqueous solution at 30 °C. The 10-dimensional hypermatrix is capable of providing all 45 possible pairwise interactions between adjacent amino acids; for simplicity, we show only the correlations between NH(*i*) and N(*i*−1) on the preceding residue. This requires a combination of a three-dimensional HNCA spectrum ( $F_1$ – $F_2$ – $F_3$ ) linking N(*i*−1) with Cα(*i*−1) and a separate three-dimensional HN(CO)CA spectrum ( $F_6$ – $F_7$ – $F_3$ ) that links NH(*i*) with the same Cα site. A virtual five-dimensional ( $F_1$ – $F_2$ – $F_3$ – $F_7$ – $F_6$ ) subspectrum is thus created, and this is then projected onto the  $F_6$ – $F_2$  plane to give the desired correlation (Figure 5.6).

## RELATED ARTICLES IN THE ENCYCLOPEDIA OF MAGNETIC RESONANCE

### Protein Dynamics from NMR Relaxation

## REFERENCES

1. W. P. Aue, E. Bartholdi, and R. R. Ernst, *J. Chem. Phys.*, 1976, **64**, 2229–2246.
2. R. N. Bracewell, ‘*Fourier Analysis and Imaging*’, Kluwer Academic: New York, 2004.
3. G. Bodenhausen and R. R. Ernst, *J. Am. Chem. Soc.*, 1982, **104**, 1304–1309.
4. Ě. Kupče and R. Freeman, *J. Am. Chem. Soc.*, 2004, **126**, 6429–6440.
5. K. Ding and A. M. Gronenborn, *J. Magn. Reson.*, 2002, **156**, 262–268.
6. S. Kim and T. Szyperski, *J. Am. Chem. Soc.*, 2003, **125**, 1385–1393.
7. W. Kozminski and I. Zhukov, *J. Biomol. NMR*, 2003, **26**, 157–166.
8. Ě. Kupče and R. Freeman, *J. Magn. Reson.*, 2008, **191**, 164–168.
9. Ě. Kupče and R. Freeman, *Prog. NMR Spectrosc.*, 2008, **52**, 22–30.

