

See discussions, stats, and author profiles for this publication at:  
<https://www.researchgate.net/publication/282642302>

# Covariance NMR and Small Molecule Applications

ARTICLE *in* ANNUAL REPORTS ON NMR SPECTROSCOPY · JULY 2014

Impact Factor: 2.27 · DOI: 10.1016/B978-0-12-800183-7.00005-8

---

CITATION

1

READS

29

## 2 AUTHORS:



Martin Jaeger

University Hochschule Nied...

24 PUBLICATIONS 125

CITATIONS

SEE PROFILE



Ruud Aspers

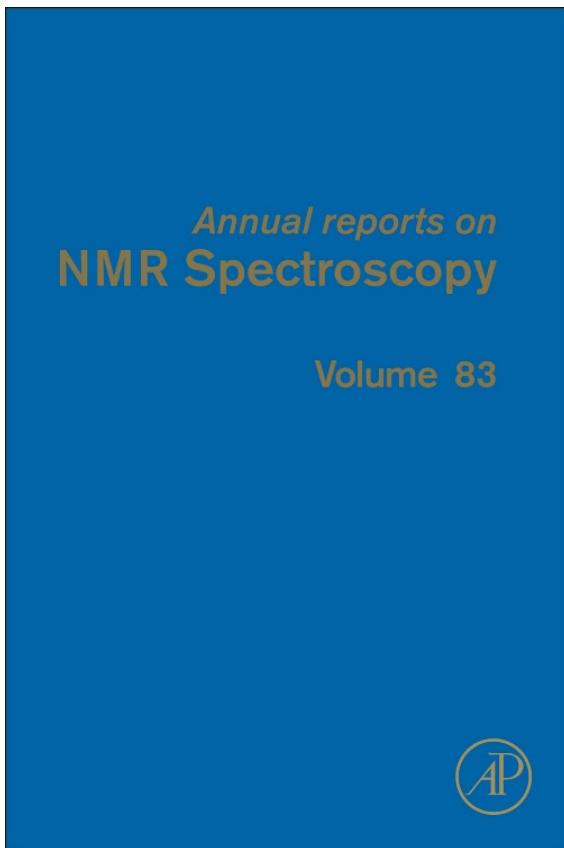
Radboud University Nijmegen

14 PUBLICATIONS 56 CITATIONS

SEE PROFILE

**Provided for non-commercial research and educational use only.  
Not for reproduction, distribution or commercial use.**

This chapter was originally published in the book *Annual Reports on NMR Spectroscopy*, Vol. 83 published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who know you, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<http://www.elsevier.com/locate/permissionusematerial>

From: Martin Jaeger, Ruud L.E.G. Aspers, Covariance NMR and Small Molecule Applications. In Graham A. Webb, editor: *Annual Reports on NMR Spectroscopy*, Vol. 83, Oxford: United Kingdom, 2014, pp. 271-349.  
ISBN: 978-0-12-800183-7  
© Copyright 2014 Elsevier Ltd.  
Academic Press



# Covariance NMR and Small Molecule Applications

**Martin Jaeger<sup>\*,†</sup>, Ruud L.E.G. Aspers<sup>‡</sup>**

<sup>\*</sup>Department of Chemistry, Instrumental Analytical Chemistry, Niederrhein University of Applied Sciences, Krefeld, Germany

<sup>†</sup>ILOC Institute for Coatings and Surface Chemistry, Niederrhein University of Applied Sciences, Krefeld, Germany

<sup>‡</sup>Institute for Molecules and Materials, Biophysical Chemistry, Radboud University, Nijmegen, The Netherlands

Dedicated to Professor Sybren S. Wijmenga, NMR connoisseur and friend.

## Contents

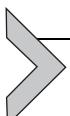
1. Introduction	272
2. On the Theory of Covariance NMR	273
2.1 The General Description	273
2.2 Types of Covariance Transformations in NMR	278
2.3 Examples of Covariance Transformations	284
2.4 Artefacts in Covariance NMR Spectra	289
2.5 The Determination of the Non-Linear Signal-to-Noise Ratio	290
2.6 Asynchronous Spectra—The Neglected Imaginary Part	291
2.7 Heterospectroscopy	294
3. Software for Covariance NMR Processing	296
4. Covariance and NUS—The Combination of Two Approaches to Fast Methods	300
5. Applications of Covariance NMR to Small Molecules	306
5.1 Direct Covariance	306
5.2 Indirect Covariance	310
5.3 Unsymmetrical and Generalized Indirect Covariance	318
5.4 Multidimensional Covariance	333
5.5 Synchronous and Asynchronous Spectra	335
5.6 Statistical Analysis of NMR Spectra in the Field of Metabolomics	336
5.7 Other Covariance Applications in NMR Spectroscopy	338
6. Conclusion	340
Acknowledgements	341
References	341

## Abstract

Covariance NMR or the processing of NMR data according to statistical principles has grown into a powerful tool for structure elucidation, signal assignments and identification of mixture constituents. Experimental data processing by covariance can either

replace or accompany the traditional Fourier transformation. The covariance formalism for NMR spectroscopy is visualized through matrix representations of data arrays or spectra and their algebraic transformations. Illustrative examples are given as simple matrix calculations for the most common covariance processing strategies. Four types of covariance treatments are distinguished for use with 2-dimensional NMR: direct, indirect, unsymmetrical or generalized, and multidimensional covariance. The covariance formalism proved its strength in enhancing spectral resolution, increasing sensitivity or decreasing experiment time, and providing access to spectra displaying correlations between insensitive heteronuclei. In particular, the combination with non-uniform sampling and pure shift or homo-decoupling techniques is discussed. The use of synchronous and asynchronous correlation spectra, as known from other spectroscopic technologies, is reviewed with respect to NMR. Applications to small molecules are collected, providing an introduction to the interpretation and demonstrating the practical aspects, the results and the advantages of the covariance processing strategies, such as improved sensitivity and resolution or the visualization of exotic correlations.

**Keywords:** Covariance NMR, Synchronous and asynchronous spectra, Non-uniform sampling, Fast methods, Structure elucidation



## 1. INTRODUCTION

The quest for the enhancement of sensitivity, specificity and resolution is ubiquitous in all analytical and spectroscopic sciences. In the tradition of NMR spectroscopy, the creation of stronger magnets, thus higher magnetic field strengths used to provide the desired increase in sensitivity and resolution. The development of pulsed-field gradients and the advent of cryogenically cooled probes accounted for another manner to augment instrumental sensitivity. In the more recent past, so-called fast NMR methods attracted vivid interest to push the limits of sensitivity even further. These techniques followed different paths towards improved sensitivity. Hardware-oriented strategies, such as parallel acquisition and detection, pulse sequence-based methods, such as time-shared experiments, and the combination of several experiments into one pulse sequence, were devised. The optimal use of the available experimental time within a defined experiment shifted into focus when data sampling schemes were optimized. In this approach, the common linear acquisition sampling was replaced by non-linear, statistical or sparse sampling methods, especially for multi-dimensional experiments. The spectral processing procedure was equally subjected to thorough investigation. In this respect, one of the vital organs of modern pulse NMR, the Fourier Transform, was dissected. Statistic data treatment allowed the transformation of experimentally acquired data into spectra equally well. Derived from the statistical mathematical

theory, the methodology became known as covariance NMR or covariance processing. In its most general form, it is referred to as covariance spectroscopy.

Due to its purely mathematical nature, the speed of covariance NMR processing only depends on the computational power available. In terms of sensitivity and resolution it reflects the properties of the spectrum or spectra, in particular, the acquisition dimension, to be subjected to covariance treatment. More details will follow in the course of this review.

The development and thorough investigation of covariance NMR spectroscopy is intimately connected to Brüschweiler and coworkers, and Martin and coworkers. From the former group's contribution, the major theoretical foundations and single-experiment transformations might be explicitly mentioned. Among the latter group's achievements in covariance NMR, the thorough artefact analysis and the transformations of component spectra to hyphenated-experiment spectra might be emphasized at this point. In the beginning of covariance NMR some 10–13 years back, Noda and coworkers had already introduced covariance transformations within the field of infrared spectroscopy (IR or FTIR). They designed and explored general covariance transformations in that field, starting out with 2-dimensional IR, where the correlation was achieved via a common perturbation domain, for example, a temperature variation, a chemical reaction or other sample variations. Eventually, they approached covariance NMR as well.

The achievements and discoveries of the last decade have been summarized in several reviews [1–4]. The following section is intended to give an introductory reading into the theory of covariance NMR and to provide an overview of the applications to small and—to some extent—larger molecules.



## 2. ON THE THEORY OF COVARIANCE NMR

### 2.1. The General Description

Covariance NMR mostly refers to any NMR experiment whose resulting data are subjected at some point to covariance analysis, covariance transformation or covariance treatment. Covariance NMR processing describes the steps that compute the covariance from a matrix of NMR data and yields the covariance map. The covariance map is equivalent to a NMR spectrum obtained after Fourier transformation, if the covariance was calculated obeying certain mathematical constraints, cf. further below. In other words,

covariance NMR processing refers to the generation of an NMR spectral representation where the correlations between magnetic nuclear spins are obtained from the quantitative comparison of spectral intensity variations observed over an instrumental time interval. The covariance transformation either complements or replaces the common Fourier transformation. It may be executed along the direct or the indirect dimension. It can be applied to homonuclear or heteronuclear data sets or it can combine two or more data sets.

The term covariance as used here has its origins in statistical mathematics. Variances are used to describe the deviation from the mean of a series of data. The covariance according to Eq. (5.1) can hence be interpreted as the difference between the correlated and the uncorrelated products of a series of data.

$$\mathbf{C}(x, y) = \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle = \langle xy \rangle - \langle x \rangle \langle y \rangle \quad (5.1)$$

where  $\langle x \rangle$  and  $\langle y \rangle$  denote the mean values, and  $\langle \cdot \rangle$  represents any type of correlation function. Then,  $\mathbf{C}$  is the covariance matrix.

In 1989, Frasinski *et al.* [5] described the potential of Eq. (5.1) to generate physically meaningful spectral correlations in laser spectroscopic experiments. Equally in terms of spectroscopy, Noda [6–8] introduced the covariance calculations into the analysis of IR, inspired by the 2-dimensional correlation spectroscopic techniques known for FT 2D NMR.

Assuming that  $x$  and  $y$  of Eq. (5.1) are spectroscopic data series and that they are arranged such that  $S$  is a spectrum of  $N_1$  data points, the elements  $C_{ij}$  of the covariance matrix  $\mathbf{C}$ , also called the covariance map, are computed according to Eq. (5.2)

$$C_{ij} = \frac{1}{N_1 - 1} \sum_{k=1}^{N_1} (S(k, i) - \langle S(i) \rangle)(S(k, j) - \langle S(j) \rangle) \quad (5.2)$$

where  $\langle S(i) \rangle$  is the average spectrum as defined in Eq. (5.3).

$$\langle S(i) \rangle = \frac{1}{N_1} \sum_{k=1}^{N_1} S(k, i) \quad (5.3)$$

Substitution of  $i$  by  $j$  leads to the equivalent definition for  $\langle S(j) \rangle$ .

While for mathematical purposes Eqs. (5.1) and (5.2) are appropriate, the symbols for time and frequency, that is,  $t$  and  $\nu$  or  $\omega$ , are more common for

spectroscopic considerations. By means of Parseval's [theorem \(5.4\)](#), Eqs. [\(5.1\)](#) and [\(5.2\)](#) can be transformed such that the covariance matrix can be expressed as Eq. [\(5.5\)](#).

$$\int_{-\infty}^{\infty} f(t)g^*(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)G^*(\omega)d\omega \quad (5.4)$$

The variables  $T_{\max}$  and  $T_{\min}$  represent the limits of the time interval the second or indirect dimension is recorded in. The index inc refers to the indirect spectral dimension. As will later be seen, the direct dimension can originate from two different data sets, then  $A \neq B$ , or from the same data set, hence  $A = B$ . In the latter case, one data set is merely the transpose of the other one. Within a representation such as Eq. [\(5.5\)](#), the two data sets—which contain mixed time–frequency data before and frequency–frequency data after transformation—are correlated by a shared indirect dimension. The common feature can be understood as a perturbation, the dimension hence called perturbation dimension.

$$\begin{aligned} C(\omega_{2,A}, \omega_{2,B}) &= \langle s(t_{\text{inc}}, \omega_{2,A}) \cdot s(t_{\text{inc}}, \omega_{2,B}) \rangle \\ &= \frac{1}{2\pi(T_{\max} - T_{\min})} \int_0^{\infty} S(\omega_{\text{inc}}, \omega_{2,A}) \cdot S^*(\omega_{\text{inc}}, \omega_{2,B}) d\omega_{\text{inc}} \quad (5.5) \\ &= \Phi(\omega_{2,A}, \omega_{2,B}) + i\Psi(\omega_{2,A}, \omega_{2,B}) \end{aligned}$$

where  $\Phi$  and  $\Psi$  are defined according to Eqs. [\(5.6\)](#) and [\(5.7\)](#).

$$\Phi(\omega_{2,A}, \omega_{2,B}) = \frac{1}{T_{\max} - T_{\min}} \int_{T_{\max}}^{T_{\min}} s(t_{\text{inc}}, \omega_{2,A}) \cdot s(t_{\text{inc}}, \omega_{2,B}) dt_{\text{inc}} \quad (5.6)$$

$$\Psi(\omega_{2,A}, \omega_{2,B}) = \frac{1}{T_{\max} - T_{\min}} \int_{T_{\max}}^{T_{\min}} s(t_{\text{inc}}, \omega_{2,A}) \cdot h \cdot s(t_{\text{inc}}, \omega_{2,B}) dt_{\text{inc}} \quad (5.7)$$

where  $h$  is the Noda–Hilbert transform, cf. also Eqs. [\(5.17\)](#) and [\(5.18\)](#) for definition and matrix notation.

It is instructive to inspect the integrated forms of Eqs. [\(5.6\)](#) and [\(5.7\)](#) as given in Eqs. [\(5.8\)](#) and [\(5.9\)](#).

$$\Phi(\omega_{2,A}, \omega_{2,B}) = p(\cos\varphi)^{A,B} \text{Abs}(\omega_{2,A}) \text{Abs}(\omega_{2,B}) \quad (5.8)$$

$$\Psi(\omega_{2,A}, \omega_{2,B}) = q(\sin\varphi)^{A,B} \text{Abs}(\omega_{2,A}) \text{Abs}(\omega_{2,B}^*) \quad (5.9)$$

The complete representations of Eqs. (5.8) and (5.9) result in lengthy expressions for  $p$  and  $q$  containing summation and integration terms. Therein, a phase  $\varphi$  can be recognized that may be interpreted in terms of an internal reference, according to Eqs. (5.10) and (5.11), which show the important factors from the complete definition of  $p$  and  $q$ .

$$p(\cos \varphi)^{A,B} \cos(\omega_{2,\alpha} t_{\text{inc}} + \varphi), \quad \alpha = A, B \quad (5.10)$$

$$q(\sin \varphi)^{A,B} \sin(\omega_{2,\alpha} t_{\text{inc}} + \varphi), \quad \alpha = A, B \quad (5.11)$$

When comparing the analogous expression in a representation after Fourier transformation, cf. Eq. (5.12), it can easily be recognized that an internal reference like  $\varphi$  is absent, that is, manual phase correction after FT is needed.

$$\begin{aligned} S(\omega_{\text{inc}}, \omega_2) &= \int s(\omega_{\text{inc}}, \omega_2) \exp(-i\omega_2 t_{\text{inc}}) dt_{\text{inc}} \\ &= \int s(t_{\text{inc}}, \omega_2) \cos(\omega_{\text{inc}} t_{\text{inc}}) dt_{\text{inc}} + i \int s(t_{\text{inc}}, \omega_2) \sin(\omega_{\text{inc}} t_{\text{inc}}) dt_{\text{inc}} \end{aligned} \quad (5.12)$$

It should be noted that the spectrum after FT is often only the real part. Yet, the phasing reference consideration remains the same. The interested reader is referred to classical NMR textbooks and to the recent works in the context of covariance NMR cited in this review [9,10].

From Eqs. (5.1) and (5.2), two conclusions can quickly be drawn. Firstly, Eq. (5.1) can be extended to Eq. (5.13), suggesting directly that covariance transformations can be amended to three and higher dimension.

$$\mathbf{C}(x, y, z) = \langle (x - \langle x \rangle)(y - \langle y \rangle)(z - \langle z \rangle) \rangle \quad (5.13)$$

The generalization of Eq. (5.3) is given as Eq. (5.14). The variables  $f$  and  $\omega$  denote spectral variables such as frequencies, that may be obtained from any kind of spectroscopy. They are only related by the common time domain  $t$ , which can also be substituted by another perturbation dimension, such as a series of samples.

$$\mathbf{C}(f, \omega) = \langle s_1(f, t) \cdot s_2(\omega, t) \rangle \quad (5.14)$$

According to Eq. (5.14), covariance transformations can be applied to any type of spectral data sets that are connected to each other by a common history or domain. The spectra thus generated represent heterospectral correlation maps. In NMR spectroscopy, this concept was taken up as unsymmetrical indirect covariance (UIC) NMR, relating, for example,

$^{15}\text{N}$  and  $^{13}\text{C}$  signals via the proton dimension to each other. Even more generalized, NMR and IR or NMR and mass spectrometry (MS) data were correlated.

The elements  $C_{ij}$ , which are the mathematical covariances between the amplitudes of positions  $i$  and  $j$  of the 1D spectra, form the covariance matrix  $\mathbf{C}$ . Based on Eq. (5.2),  $\mathbf{C}$  can be expressed as the matrix product of  $\mathbf{S}$ , which is the spectroscopic data set. The matrix multiplication of  $\mathbf{S}$  with its transpose  $\mathbf{S}^T$  yields the  $\mathbf{C}^2$  as displayed in Eq. (5.15).

$$\mathbf{C}^2 = \mathbf{S}^T \cdot \mathbf{S} \quad (5.15)$$

Since thorough mathematical derivations and considerations have been described by Brüschweiler *et al.* [11] and Noda *et al.* [12], only three aspects shall be emphasized here.

Firstly, using the definition of  $\mathbf{S}$  as being the mixed time–frequency matrix,  $\mathbf{S}(t_1, \omega_2)$ , the product  $\mathbf{S}^T \mathbf{S}$  yields a symmetric matrix  $\mathbf{C}(\omega_1, \omega_2)$ . Secondly, assuming in good spectroscopic practise,  $\mathbf{S}$  has dimensions of  $N_1 = 2\text{ k}$  and  $N_2 = 256\text{ k}$  data points, the resulting covariance map will represent a spectrum of  $N_1 \times N_2 = 2\text{ k} \times 2\text{ k}$  data points. This transformation can be interpreted in terms of projecting the direct or acquisition dimension onto the indirect or incremented dimension. As a consequence, the resolution of the indirect dimension is increased to the level of the direct dimension. Thirdly, the covariance map can also be generated from two data matrices  $\mathbf{F}^T$  and  $\mathbf{F}$  that have been Fourier transformed in both dimensions. The mathematical expression is given in Eq. (5.16).

$$\mathbf{C}^2 = \mathbf{S}^T \cdot \mathbf{S} = \mathbf{F}^T \cdot \mathbf{F} \quad (5.16)$$

The equality of transformations of the mixed time–frequency data and the completely Fourier transformed data is a consequence of Parseval’s theorem (5.4) and has been previously discussed [10,11]. It can also be understood by taking into account that the spectral reconstruction is achieved by relating two direct dimensions via an indirect dimension, which in turn is discarded. Whether two frequency domains are correlated via a common time domain or a common frequency domain is therefore equivalent. It is also equivalent to Noda’s model of relating two IR wavenumber dimensions via a common perturbation stemming from either time or sample space. From the matrix representation, it can be seen that Noda’s synchronous matrix  $\Phi$ , Eqs. (5.6) and (5.17) corresponds to the covariance map according to Eq. (5.16), if the data matrices yielding  $\Phi$  are composed of

mean centred spectra. As will be shown later, this calculation usually relates two direct or observed dimensions, which was hence called direct covariance. The asynchronous map  $\Psi$  of Eqs. (5.7) and (5.18) corresponds to the indirect covariance correlation plot. Furthermore, the Eqs. (5.16) and (5.14) implicitly open up the field of covariance spectroscopy to the application of heterocorrelation spectroscopy.

The matrix notation of Eqs. (5.6) and (5.7) are given here as Eqs. (5.17) and (5.18).

$$\Phi = \bar{\mathbf{X}}^T \cdot \bar{\mathbf{X}} \quad (5.17)$$

$$\Psi = \bar{\mathbf{X}}^T \cdot \mathbf{N} \cdot \bar{\mathbf{X}} \quad (5.18)$$

where  $\bar{\mathbf{X}}$  is the matrix of mean centred spectra and  $\mathbf{N}$  the Noda–Hilbert orthogonalization matrix with  $N_{ik}$  equals 0 if  $i=k$  and  $1/(\pi(k-i))$  otherwise.

## 2.2. Types of Covariance Transformations in NMR

### 2.2.1 The Classification of Covariance NMR

As a consequence of Eq. (5.16), four covariance NMR types can be defined. The use of symbols is the same as above:  $\mathbf{C}$  is the covariance matrix, that is, the spectrum,  $\mathbf{S}$  is the mixed time–frequency 2D data representation,  $\mathbf{F}$  and  $\mathbf{G}$  are frequency–frequency data matrices, that is, spectra;  $\mathbf{G}_{\text{reg}}$  is a regularized data matrix according to Eq. (5.26).

#### 1. Direct covariance

$$\mathbf{C} = (\mathbf{S}^T \cdot \mathbf{S})^{1/2} = (\mathbf{F}^T \cdot \mathbf{F})^{1/2} \quad (5.19)$$

where the same spectrum is subjected to the covariance transformation (5.19) such that the direct or acquisition domain is preserved. A symmetric covariance map, that is equivalent to the 2D FT NMR spectrum but with enhanced resolution in the indirect dimension, is obtained.

#### 2. (a) Indirect covariance

$$\mathbf{C} = (\mathbf{F} \cdot \mathbf{F}^T)^{1/2} \quad (5.20)$$

where the same spectrum is subjected to covariance treatment according to Eq. (5.20) such that the indirect dimension is preserved. A symmetric spectrum correlating signals in the indirect dimension is obtained.

**(b)** Doubly indirect covariance

$$\mathbf{C} = \mathbf{F} \cdot \mathbf{G}_{\text{reg}} \cdot \mathbf{F}^T \quad (5.21)$$

where the spectrum resulting from Eq. (5.21) is similar to an indirect covariance spectrum [13]

**3. (a)** Unsymmetrical indirect covariance

$$\mathbf{C} = \mathbf{F} \cdot \mathbf{G}^T \quad (5.22)$$

where two 2D spectra  $\mathbf{F}$  and  $\mathbf{G}$  are subjected to covariance processing following Eq. (5.22) such that a spectrum correlating both indirect dimensions results.

As can be seen, UIC according to Eq. (5.22) does not lead to a symmetric matrix, unless  $\mathbf{F}$  and  $\mathbf{G}$  are themselves symmetric. Therefore, the square root operation must not be applied. It has been discussed that the lack of this operation causes a more significant presence of artefacts from the mathematical operations. As a remedy, generalized indirect covariance (GIC) according to Eq. (5.23) was proposed [14].

**(b)** Generalized indirect covariance

$$\mathbf{C} = \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{F}^T & \mathbf{G}^T \end{bmatrix} = \begin{bmatrix} \mathbf{F} \cdot \mathbf{F}^T & \mathbf{F} \cdot \mathbf{G}^T \\ \mathbf{G} \cdot \mathbf{F}^T & \mathbf{G} \cdot \mathbf{G}^T \end{bmatrix} \quad (5.23)$$

where one of the off-diagonal elements of the GIC matrix equals the UIC map.

Through the application of single value decomposition (SVD) of the symmetric matrix, the  $\lambda$ th power and hence the square root of the UIC spectrum is defined via the GIC matrix. Spectral artefacts due to covariance processing can be distinguished from native signals on account of their dependence on  $\lambda$ . Further details on the GIC matrix calculation are given below, cf. Eq. (5.27).

**4. Multidimensional covariance**

**(a)** Triple-rank covariance

$$\mathbf{H}_{ijkl\dots} = \mathbf{D}_{ij} \mathbf{E}_{kj} \mathbf{F}_{kl} \mathbf{G}_{ml} \dots \quad (5.24)$$

where  $\mathbf{D}, \mathbf{E}, \mathbf{F}, \mathbf{G}, \dots$  are standard 2D FT NMR spectra, each consisting of  $N_1 \times N_2$  data points. Indices  $i, k, m$  range from 1 to  $N_1$  and  $j, l$  from 1 to  $N_2$ . A filter function, for example, based on spectral moments, can be applied for spectral enhancement. Summation

over all indices that occur twice in Eq. (5.20) affords the covariance matrix, equivalent to UIC processing:  $\mathbf{C} = \mathbf{D} \mathbf{E}^T \mathbf{F} \mathbf{G}^T \dots$

The result of a 3R transformation according to Eq. (5.24) is a 3D representation built from the parent 2D FT NMR spectra. The triple-rank (3R) covariance formalism was derived by Bingol *et al.* [15]

**(b)** 4-dimensional covariance

$${}^2C_{ij} = \text{Tr} \left( \left( \mathbf{D}_{\omega'_3 \omega'_4} \right)^T \cdot \mathbf{D}_{\omega''_3 \omega''_4} \right) \quad (5.25)$$

where  $\text{Tr}$  is the trace of a matrix. The 4D spectrum  $\mathbf{F}(\omega_1, \omega_2, \omega_3, \omega_4)$  can be represented as an array of 2D data sets composed of  $\mathbf{A}_{\omega_1, \omega_2}$  and  $\mathbf{D}_{\omega_3, \omega_4}$ . The elements  ${}^2C_{ij}$  of the covariance matrix are calculated as the inner product between the donor planes  $\mathbf{D}_{\omega'_3, \omega'_4}$  and  $\mathbf{D}_{\omega''_3, \omega''_4}$  associated with the acceptor pairs  $(\omega'_3, \omega'_4)$  and  $(\omega''_3, \omega''_4)$ . Equation (5.25) thus defines the covariance signal at the position  $(\omega'_3, \omega'_4, \omega''_3, \omega''_4)$ . The full description of 4D covariance is given in Ref. [16].

### 2.2.2 Some Aspects of the Workings of 4D NMR

The potential to extend 2-dimensional covariance NMR to higher dimensionality has its foundations in Eq. (5.13). Thus, Snyder *et al.* [16] laid the basis for the computation of 4D NOESY spectra. In their strategy, the critical entry point consisted of considering a 4D data set an array of 2D data or a plane-of-planes. In order to illustrate the calculations, the terms donor and acceptor planes in combination with donor and acceptor pairs were coined. It should be noted that an acceptor plane is associated with each donor pair  $(\omega_1, \omega_2)$  at frequencies  $\omega_1$  and  $\omega_2$ . ‘Mapping’ of either the acceptor planes onto the donor planes or vice versa describes the projection of a dimension onto another, which leads to an increase for direct covariance or to a decrease for indirect covariance in dimensionality of resolution.

A spectrum  $\mathbf{F}(\omega_1, \omega_2, \omega_3, \omega_4)$  can be decomposed into  $\mathbf{A}_{\omega_1, \omega_2} = \mathbf{F}(\omega_1, \omega_2, \cdot, \cdot)$  and  $\mathbf{D}_{\omega_3, \omega_4} = \mathbf{F}(\cdot, \cdot, \omega_3, \omega_4)$ , where  $\omega_i$  denote fixed chemical shift values. Matrices  $\mathbf{A}$  and  $\mathbf{D}$  are referred to as acceptor and donor planes, respectively. The matrix is set up such that the acceptor planes are composed with the higher resolution of the direct or acquisition dimension. The replacement of the donor planes with the acceptor planes is said to yield

an acceptor projection. Prior to covariance calculation, a threshold filter may be applied, thus reducing computational time.

The covariance map is calculated as the inner product between donor planes,  $\mathbf{D}_{\omega'_3, \omega'_4}$  and  $\mathbf{D}_{\omega''_3, \omega''_4}$  according to Eq. (5.25). The 2D representation of the 4D covariance spectrum is the square root of  $\mathbf{C}^2$ , whose elements  ${}^2C_{ij}$  contain the acceptor pairs,  $(\omega'_3, \omega'_4)$  and  $(\omega''_3, \omega''_4)$ . The inner product of the two donor planes hence constitutes the covariance signal at the position  $(\omega'_3, \omega'_4, \omega''_3, \omega''_4)$ .

### 2.2.3 Matrix Regularization in Covariance NMR

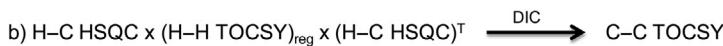
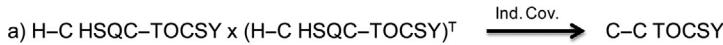
The regularized covariance map is defined according to Eq. (5.26)

$$\mathbf{G}_{\text{reg}} = \text{Abs}\left(\left(\mathbf{G}_\alpha^T \cdot \mathbf{G}_\alpha\right)^{1/2} - \alpha \cdot \mathbf{1}\right) \quad (5.26)$$

where  $\text{Abs}$  denotes the absolute value of each matrix element. The matrix  $\mathbf{G}_\alpha$  equals  $\mathbf{G} + \alpha \mathbf{1}$  with the unity matrix  $\mathbf{1}$  and the regularization factor  $\alpha$ . The regularization procedure allows the introduction of scaled diagonal peaks into a covariance map when they were suppressed by the pulse sequence. The presence of strong diagonal elements is required if the correlations signals were to be used in a quantitative manner [17]. According to the discrete nature of the data, scaling as well as filtering can be performed differently for individual spectral regions.

### 2.2.4 The Transition from Indirect to Unsymmetrical Indirect Covariance

Doubly indirect covariance (DIC) yields a symmetric correlation map [13]. It can further be considered as a bridge to the generalization of indirect covariance. It implies that correlation spectra, that are used for indirect covariance, themselves may be a creation of a previous covariance transformation. Scheme 5.1 serves as an example to elucidate this perspective. In Scheme 5.1A and B, indirect covariance generates homonuclear spectra. Part (C) leads to the creation of a heteronuclear correlation, which itself can be taken as component spectrum for transformation (A) or (B). It is apparent that Scheme 5.1 represents an oversimplification of the types of indirect covariance treatment. A wider variety of spectra that can be obtained by UIC processing is collected in Table 5.1.



**Scheme 5.1** Indirect covariance is sketched in (A), doubly indirect covariance (DIC) in (B) and unsymmetrical indirect covariance in (C). The powers of  $1/2$  and regularization factors have been omitted for simplicity. Common experimental abbreviations were used equivalent to the spectra or data matrices obtained from the corresponding experiment.

### 2.2.5 The Workings of Generalized Indirect Covariance NMR

The relation between the GIC matrix and the spectral data of two different spectra **F** and **G** was described by Eq. (5.22). The full discussion of the derivation can be found in Ref. [14]. A brief description of the key elements follows. Let **S** be a stack of 2D spectra **X**<sub>1</sub> … **X**<sub>n</sub>, then **S** can be expressed by Eq. (5.27).

$$\mathbf{S} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_n \end{bmatrix} \quad (5.27)$$

The generalized covariance matrix shall be defined according to Eq. (5.28).

$$\begin{aligned} \mathbf{C} = \mathbf{S} \cdot \mathbf{S}^T &= \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_n \end{bmatrix} \cdot [\mathbf{X}_1^T \ \dots \ \mathbf{X}_n^T] \\ &= \begin{bmatrix} \mathbf{X}_1 \cdot \mathbf{X}_1^T & \dots & \mathbf{X}_1 \cdot \mathbf{X}_n^T \\ \vdots & \ddots & \vdots \\ \mathbf{X}_n \cdot \mathbf{X}_1^T & \dots & \mathbf{X}_n \cdot \mathbf{X}_n^T \end{bmatrix} \end{aligned} \quad (5.28)$$

SVD of matrix **S** follows Eq. (5.29), where matrix **D** is diagonal and contains the eigenvalues of the GIC matrix. Matrices **U** and **V** are orthogonal and consist of the eigenvectors of the GIC matrix.

$$\mathbf{S} = \mathbf{U} \cdot \mathbf{D} \cdot \mathbf{V}^T \quad (5.29)$$

**Table 5.1** Types of spectra that are obtained after unsymmetrical indirect covariance processing of base spectra

<b>1st component spectrum</b>	<b>2nd component spectrum</b>	<b>Spectrum resulting from the component spectra after covariance processing</b>
HSQC magnitude	COSY magnitude or TOCSY magnitude	HSQC–TOCSY magnitude [18,19]
HSQC edited or IDR	COSY or TOCSY magnitude	HSQC–TOCSY edited with distinction CH, CH <sub>2</sub> , CH <sub>3</sub> [20]
HSQC–TOCSY magnitude		CH <sub>m</sub> –CH <sub>n</sub> correlation, C–C COSY (ADEQUATE) [18,21]
IDR HSQC–TOCSY or edited [22]		C–C COSY with discrimination between CH, CH <sub>2</sub> , CH <sub>3</sub> [22]
HSQC	NOESY phase-sensitive	HSQC–NOESY with discrimination of direct CH <sub>n</sub> correlations from CH <sub>n</sub> NOE correlations [23]
HSQC edited	NOESY phase-sensitive	HSQC–NOESY mixed phase, no additional information
HSQC–NOESY		C–C NOE correlations [24]
IDR HSQC–TOCSY	HMBC	IDR HSQC–TOCSY–HMBC long-range HSQC–TOCSY [25]
TOCSY	HMBC	HMBC–TOCSY [14]
HSQC	HMBC	C–C correlation, C–C COSY with short- and long-range couplings [18]
HSQC	H–P HMBC	P–C correlation with long-range couplings [26]
HSQC	H–N HMBC	C–N HSQC–HMBC [27]
HSQC edited	1,1-ADEQUATE	HSQC-1,1-ADEQUATE C–C correlations with <sup>2</sup> J <sub>CH</sub> specificity [28]
HSQC	1, <i>n</i> -ADEQUATE	HSQC-1, <i>n</i> -ADEQUATE C–C correlations with <sup>3</sup> J <sub>CH</sub> and <sup>4</sup> J <sub>CH</sub> specificity [29]
HMBC	1,1-ADEQUATE	HMBC-1,1-ADEQUATE equivalent to <i>n</i> ,1-ADEQUATE but with single-quantum representation [30]
HSQC	<sup>1</sup> J <sub>CC</sub> -edited 1, <i>n</i> -ADEQUATE	<sup>1</sup> J <sub>CC</sub> -HSQC-1, <i>n</i> -ADEQUATE [31]

The use of phase-sensitive spectra as component spectra is considered.

Equation (5.28) can then be rewritten as Eq. (5.30) by substitution of Eq. (5.29).

$$\mathbf{C} = (\mathbf{U} \cdot \mathbf{D} \cdot \mathbf{V}^T) \cdot (\mathbf{V} \cdot \mathbf{D} \cdot \mathbf{U}^T) = \mathbf{U} \cdot \mathbf{D}^2 \cdot \mathbf{U}^T \quad (5.30)$$

which can be generalized to yield Eq. (5.31)

$$\mathbf{C}^\lambda = \mathbf{U} \cdot \mathbf{D}^{2\lambda} \cdot \mathbf{U}^T \quad (5.31)$$

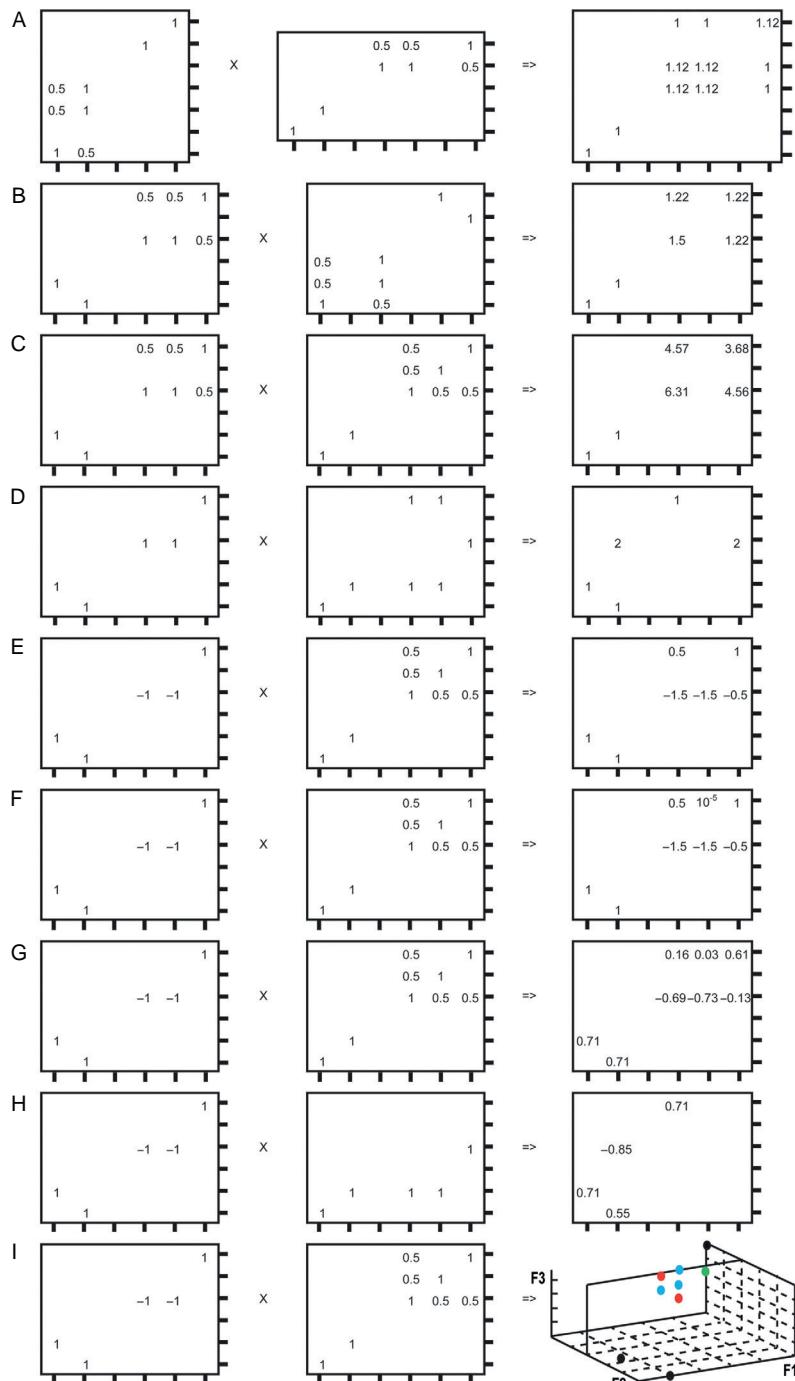
The square root operation on the covariance matrix, which is necessary to obtain the FT analogous spectrum, is calculated from Eq. (5.31) as  $\mathbf{U} \cdot \mathbf{D} \cdot \mathbf{U}^T$ , hence the square root of the diagonal eigenvalue matrix. Since the corresponding UIC transformed spectrum is contained in the GIC matrix as an off-diagonal element, the GIC formalism provides a mean to apply the square root operation on the non-symmetrical matrix  $\mathbf{F} \cdot \mathbf{G}^T$ .

### 2.3. Examples of Covariance Transformations

The matrix representation of the matrix formalism, in general, is more illustrative than the notation using sums and multiple indices. The equivalence between a 2-dimensional data matrix and a 2D spectrum or a mixed time–frequency data set is evident. While the covariance literature provides numerous examples of spectra obtained from covariance transformations of experimental data sets, which will be discussed further below, few simplified models have been designed and applied for instruction purposes [14,17,18,32–34]. Thoroughly elaborated cases are described in Refs. [6,11,12,14,35–39].

For the purpose of adding a simple model to the following discussion, Figs. 5.1 and 5.2 present an overview of the covariance transformation types 1–4 on a selected set of matrices. The arrays used are of very small dimensions as compared to experimentally recorded NMR data. Still they reflect the essential features of 2D NMR spectra such as a H–H COSY or a H–TOCSY with a short mixing time, cf. Fig. 5.1A, and an H–C HSQC, cf. Fig. 5.1B. The values of the matrix elements were arbitrarily chosen and indicate a signal or a correlation at that position. The relative signs should be interpreted as signal phases, cf. Fig. 5.1E as an example for a multiplicity-edited HSQC.

The direct covariance matrix multiplication of a homonuclear NMR spectrum yields a homonuclear covariance map whose square root is equivalent to the NMR spectrum obtained from Fourier transformation. As can be readily seen from Fig. 5.1A, the higher resolution of the direct dimension is ‘mapped’ onto the indirect dimension. Hence, a symmetric spectrum

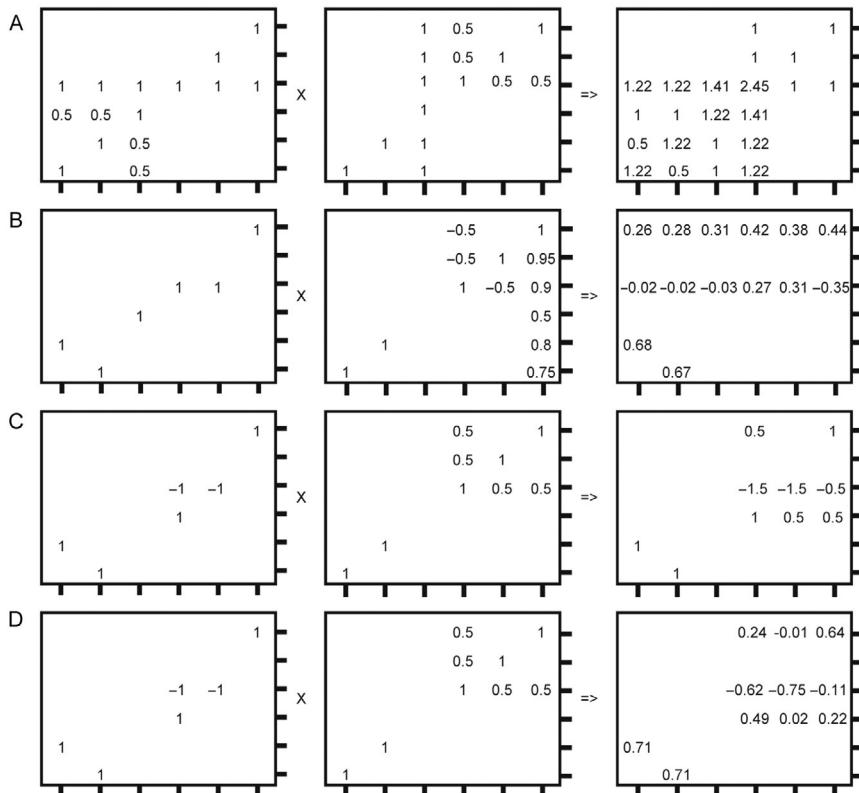


results, whose experimental indirect dimension has been enhanced with the resolution of the direct dimension. A thorough mathematical analysis by Brüschweiler [11] revealed that when applied to H–H TOCSY and H–H NOESY time–frequency data, the covariance map corresponds to its FT counterpart at another mixing time, for example, half of the mixing time in case of the TOCSY. From the example of Fig. 5.1A, it can be learned that in the limit of very low resolution, as is the case here, the correlation signals at (4,5) and (5,4) counted from the lower left corner may be either an artefact caused by the matrix formalism or a native signal liberated from the insufficient resolution of the indirect dimension.

Indirect covariance processing as depicted in Fig. 5.1B is another type of single spectrum manipulation. Here, a heteronuclear correlation spectrum is transformed into a correlation spectrum of the indirect dimension of the original data. In the current example, a H–C HSQC–TOCSY spectrum is converted into a C–C correlation spectrum, containing only correlations of carbon atoms that share the connectivity information of their attached protons. Quaternary carbons thus remain unobserved due to the principle of the H–C HSQC–TOCSY. While the direct covariance processing leads to an increase in resolution, the indirect covariance processing results in a formal decrease. Yet, the reduction of signals under retention of the information content in the case of the latter transformation compensates for the decrease in dimensionality. It is also possible to use the indirect covariance procedure on homonuclear spectra to obtain homo-decoupled

---

**Figure 5.1—Cont'd** Covariance transformations of matrices representing direct, indirect, unsymmetrical and generalized indirect, and triple-rank covariance processing of simplified 2D NMR spectra with their resulting correlation maps. Values and their signs symbolize correlation signals with intensities and phases accordingly. The matrix multiplication is indicated by a 'x'. For simplicity, no transpose operation is given in the figure. In the following legend, matrix labels are chosen according to the NMR spectra represented. (A) direct covariance of a COSY-type spectrum; (B) indirect covariance of a HSQC–COSY; (C) doubly indirect covariance of a HSQC–TOCSY and a COSY-type spectrum; (D) UIC of a HSQC and a HMBC; (E) UIC of a multiplicity-edited HSQC and COSY to yield a UIC HSQC–COSY; (F) GIC ( $\lambda=1$ ) of a multiplicity-edited HSQC and COSY to yield a GIC HSQC–COSY, the value  $10^{-5}$  may reflect numeric imperfections in the algorithm of the SVD, however a signal would be expected at this position; (G) GIC ( $\lambda=0.5$ ) of a multiplicity-edited HSQC and COSY to yield a GIC HSQC–COSY, the value of  $10^{-5}$  in the previous spectrum becomes more pronounced; (H) GIC ( $\lambda=0.5$ ) of a multiplicity-edited HSQC and a 1,1-ADEQUATE to yield the C–C correlation GIC HSQC-1,1-ADEQUATE; (I) triple-rank transformation of the multiplicity-edited HSQC and the COSY to yield a 3R spectrum, colour codes are: black circles denotes a value of 1, red (dashed open circles in print version) of  $-1$ , green (gray circles in print version) of 0.5, blue (open black circles in print version) of  $-0.5$ .



**Figure 5.2** Covariance transformations of matrices illustrating different types of artifacts. Values and their signs symbolize correlation signals with intensities and phases accordingly. The matrix multiplication is indicated by a 'x'. For simplicity, no transpose operation is given in the figure. In the following legend, matrix labels are chosen according to the NMR spectra represented. (A) direct covariance of a COSY-type spectrum (middle matrix) with a strong water resonance in the middle; (B) GIC of a HSQC and a NOESY with strong  $t_1$ -noise on the low-field resonance; (C) UIC of a multiplicity-edited HSQC and a COSY with resonance overlap at position (3,4) counted from the lower left corner; (D) GIC ( $\lambda = 0.5$ ) of the same spectra showing a reduction in signal intensity of the responses caused by the overlapping resonance.

correlation maps when decoupling in the indirect dimension was used on acquisition.

A second manner to generate the same type of correlation spectrum was introduced as doubly indirect covariance according to Eq. (5.21). The processing involves more calculations than indirect processing. The COSY matrix itself is a regularized covariance map according to Eq. (5.26). One could hence regard the DIC transformation as first generating a HSQC–COSY from a HSQC and a COSY and then subjecting it

subsequently to indirect covariance. Fig. 5.1C may serve as an example. Since the resulting C–C correlation is derived from a H–C HSQC and a H–H COSY, the sensitivity is greatly enhanced as compared to generating a C–C correlation from a H–C HSQC–COSY or H–C HSQC–TOCSY with a short mixing time. Regularization of the COSY spectrum prior to indirect covariance further contributes to an improved resolution [17].

While doubly indirect covariance processing generates a correlation map from the original indirect dimension of a spectrum **F** via a second spectrum **G**, UIC combines the indirect dimensions of two spectra, their direct dimensions serving as the link and thus being eliminated from the resulting correlation map. As is illustrated in Fig. 5.1D, UIC can be used to generate homonuclear correlations from two heteronuclear correlation spectra, such as C–C correlation maps from H–C HSQC and H–C HMBC spectra. Else, a heteronuclear H–C correlation, for example, H–C HSQC–COSY can be compiled from a H–C HSQC and a H–H COSY, cf. Fig. 5.1E. It is obvious that the covariance spectrum reflects the much higher sensitivity from the component spectra as compared to its experimental counterpart. Alternatively, it can be obtained in a much shorter experiment time [19,23,24]. The application to produce exotic correlation spectra, such as  $^{13}\text{C}$ – $^{15}\text{N}$  spectra of unlabelled compounds, that would normally suffer from inherent low sensitivity and possess hence limited accessibility, was recognized and exploited [27,40–42]. The transformation of H–C HSQC and H–H NOESY to C–C NOESY spectra by UIC was reported as well [24].

Another valuable aspect of the matrix calculation is the preservation of the sign of a matrix element, which is the spectral phase [20]. As an example may serve the UIC transformation of an edited H–C HSQC together with an H–H COSY to yield an H–C HSQC–COSY, cf. Fig. 5.1E. An overview of the use of phase-sensitive spectra in covariance processing was given by Aspers *et al.* [24]. Their collection of useful indirect covariance co-processing is extended and presented in Table 5.1.

According to the definition of the covariance matrix, its square root corresponds to the FT NMR spectrum. Since the UIC treatment, in general, leads to an unsymmetrical matrix, the square root operation is not defined. The UIC covariance map will therefore lack the full equivalence to the FT counterpart. Snyder and Brüschweiler devised the GIC formalism [14] where the UIC matrix is embedded into an array of matrices according to Eqs. (5.22) and (5.28). An illustration is given in Fig. 5.1F. The GIC array is symmetric and, by SVD, the square root and other power operations can be executed on the UIC submatrix. This is sketched in Fig. 5.1F and G for

GIC of the simplified models of the COSY and the multiplicity-edited HSQC. The spectrum resulting for  $\lambda=1$  resembles the UIC spectrum. Yet, the signal intensities are different when the square root operation  $\lambda=0.5$  is applied, see Fig. 5.1G. The positive effect of the GIC concerning the minimization of artefacts will be summarized in due course. It is not visible in this model. Fig. 5.1H symbolizes the transformation of a multiplicity-edited HSQC with an 1,1-ADEQUATE to yield a C–C correlation map, whose interpretation scheme will be given in Fig. 5.14. The construction of a 3R cube from two 2D spectra is illustrated in Fig. 5.1I. It is the only covariance transformation among the examples that does not result from matrix multiplication but from reconstruction according to Eq. (5.24).

## 2.4. Artefacts in Covariance NMR Spectra

Next to the great advantages of covariance processing, significant attention has been paid to artefacts, that is, false correlations [14,17,18,24,43–46]. In the same respect that the finite, discrete data sampling of NMR spectroscopy causes artefacts when Fourier transformed, the matrix calculations, which naturally perform more robust on discrete data, may introduce physically forbidden signals. From an NMR spectral perspective, artefacts were divided into relay effects, chemical shift degeneracy or spectral crowding in the direct and indirect dimension, and noise transposal from  $t_1$ -noise or residual water signals [1,18,21,24,47]. An overview of the origin of artefacts demonstrated on again on simple matrix models is given in Fig. 5.2. It can be readily understood from symmetry considerations that the trace of a residual water signals in a homonuclear correlation spectrum will be spread over the covariance map when direct covariance processing is applied, cf. Fig. 5.2A. Regularization, threshold filtering and spectral decomposition were suggested as remedies [17]. Especially, the removal of the data part dominated by the water signal prior to covariance processing proved successful [47]. The same water signal will disappear after UIC processing of a COSY with an HSQC data set that does not display a resonance at the water frequency. If correlations at the water signal position are present in the HSQC matrix, correlations will evidently appear in the HSQC–COSY spectrum. A similar effect is found for the transposal of  $t_1$ -noise, for example, in H–H NOESY spectra when subjected to covariance treatment, cf. Fig. 5.2B.

In contrast to the wide-spreading artefact signals, single correlation artefacts appear at specific individual frequency positions and are harder

to distinguish from native signals. Such artefacts arise from chemical shift degeneracy. Due to proton–proton signal overlap in congested or crowded spectral regions, correlations are transferred. Those false signals cannot easily be discriminated from physically legitimate correlations. The transfer due to resonance overlap, leading to false signals, has been coined relay [14,18,21]. According to their origin and their appearance in UIC C–C correlation spectra, the artefacts have been labelled type I and type II [18]. The origin of such artefact can be isochrony, for example, in H–C HSQC–TOCSY spectra, that originates from directly or remotely connected nuclei. The artefacts can be distinguished by their different phases when phase-sensitive spectra have been recorded. Aspers *et al.* [24] tried to quantify the relation between spectral congestion and the number of artefacts by introducing a spectra density defined as signals per spectral region or area. They reported that light signal overlap or few incidents of isochrony in either dimension did not hamper structure elucidation or verification. False correlations could be traced back to the corresponding overlap in the original component spectra. In contrast strong signal congestion was described as lowering the level of confidence on structure elucidation and especially *de novo* structure elucidation. The effect of resonance overlap in the direct and indirect dimension on UIC and GIC processing with  $\lambda=0.5$  is illustrated in Fig. 5.2C and D.

Snyder and Brüschweiler demonstrated that GIC provides a mean to minimize or reduce artefacts especially within UIC spectra that are contained as submatrices in the generalized covariance matrix, cf. Eq. (5.27) [14]. The influence of the exponent  $\lambda$  as defined by Eq. (5.31) was found more pronounced on false correlations than on native signals. Therefore, comparative plots of the intensities of false and legitimate signals in spectra created by using varied values of  $\lambda$  allow to distinguish between false and true correlations. The intensity of false correlations decreases more rapidly with decreasing  $\lambda$  [14]. The effect of varying values of  $\lambda$  can be recognized in Fig. 5.2C and D. On inspection of the matrix element values, the correlations originating from shift degeneracy or relays appear to decrease faster than those that can be traced back to a uniform correlation. It should however be noted that the signal intensities of covariance spectra also depend on the number of parent signals that contribute to the correlation. The signal dependence on  $\lambda$  in complex spectra might not always lead to a correct identification of artefact responses.

## 2.5. The Determination of the Non-Linear Signal-to-Noise Ratio

As a measure of sensitivity in NMR spectroscopy, the signal-to-noise ratio is a common and convenient scale. Many NMR spectroscopists are used to a

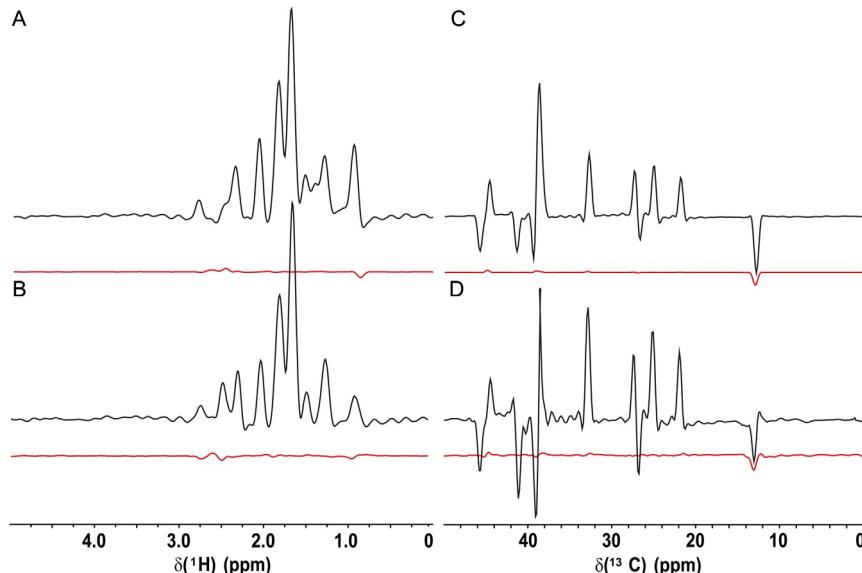
uniform noise background provided by the linearly processing Fourier transformation. Unlike FT, covariance processing is a non-linear method, thus associating the highest noise level with the most pronounced spectral change, that is, the signal, and the smoothest noise region at the largest distance from the signal. This non-uniformity causes an over- or underestimation of the signal-to-noise ratio depending on the position where the noise and the signal are measured [36]. The statistic method, known as Z-score or Z-matrix formalism, was introduced as a remedy by Snyder *et al.* [36]. The Z-formalism converts the covariance map into a Z-score map, that is, a measure of the difference in standard deviations of a statistic from its mean. The Z-score then reflects if the value of the covariance element, which is above an arbitrarily chosen threshold, stems from a signal in the input spectra or from noise. It is the thresholding that renders the noise floor homogeneous. Yet, it does not remove the noise from peaks. As part of this Z-matrix formalism, thresholding is also accompanied by maxima ratio scaling, a differential scaling procedure where each matrix element is multiplied by a weighing factor.

In summary, the signal-to-noise ratio can be estimated and compared to FT when the Z-matrix procedure according to Snyder *et al.* is used. A rough signal-to-noise estimate can still be successfully performed following the general rule of thumb: when unsymmetrical or generalized covariance processing is executed, the sensitivity of the resulting spectrum approaches that of the component spectra [21,23,36]. From visual comparison of traces from HSQC–TOCSY spectra generated by UIC and GIC from a multiplicity-edited HSQC and a TOCSY data set of the steroid tibolone, shown in Fig. 5.3, it can be readily derived that no significant abnormalities in the appearance of noise are found. The proton traces at an arbitrarily selected carbon frequency on-resonance and off-resonance, Fig. 5.3A and B, do not reveal obvious differences in their signal-to-noise ratio. Although the noise floor for on- and off-resonance traces appears slightly less flat at the signal traces, all traces are of comparable spectral quality and uniformity to allow signal identification. The same findings apply for the carbon frequency traces, Fig. 5.3C and D, although the GIC spectrum may exhibit a somewhat higher noise level. A quantitative comparison based on experimental data as well was conducted by Martin *et al.* [28].

## 2.6. Asynchronous Spectra—The Neglected Imaginary Part

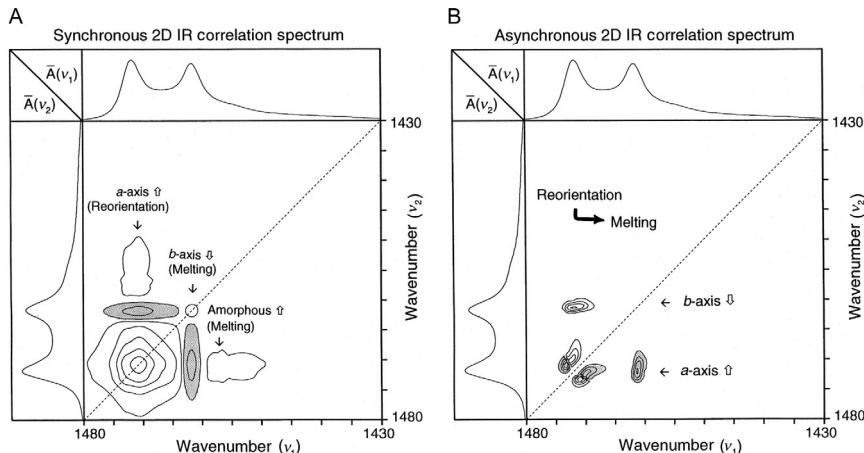
### 2.6.1 General Aspects of Asynchronous Spectra

An aspect, which has not been very intensively exploited in covariance NMR spectroscopy, is contained in Eq. (5.5). While the real part of a



**Figure 5.3**  $^1\text{H}$  (A and B) and  $^{13}\text{C}$  (C and D) 1D traces of UIC (A and C) and GIC (B and D) H-C HSQC-TOCSY spectra of tibolone **16**. Black lines represent traces containing signals; red (gray in print version) lines represent noise floors in the corresponding dimension. Tibolone spectra were considered heavily congested in both dimensions following Aspers *et al.* [24]. Residual signals can hence be recognized in the noise traces. For quantitative analyses of signal-to-noise ratios of covariance processed data, see Ref. [28].

spectral data array yields the correlations according to the type of covariance processing applied, the imaginary part gives rise to a covariance map that displays the variations of spectral intensities that proceeded successively or simultaneously during the investigated time interval. Hence, the term, synchronous and asynchronous spectra, was used by Noda [6,7]. Since Noda and coworkers started out on IR, the common dimension of 2-dimensional data arrays that were involved in the covariance formalism usually contained a macroscopic or laboratory time frame. These perturbation dimensions were realized by a reaction, solvent evaporation, biological processes, also temperature or composition variation [3,4,48]. In the asynchronous spectra according to Eq. (5.5), cross-peaks or correlations were observed when changes of spectral intensity occurred sequentially, successively or out-of-phase. The sign of a correlation signal could be used to identify whether the intensity change occurred first in the frequency domain 1 or 2 of the 2D spectral array. Four rules for the interpretation of the asynchronous map were established by Noda [48]. In order to obtain meaningful asynchronous spectra, covariance transformation has to be applied to two different



**Figure 5.4** 2D IR correlation spectra based on the spectral changes induced by the static compression of linear low-density polyethylene (LLDPE) at 608 °C. Synchronous spectrum (left): Negative intensity peaks indicating anti-correlation are shaded. Asynchronous spectrum (right): Shaded peaks indicate that intensity changes at  $\nu_1$  occur at higher pressure than changes at  $\nu_2$ . Reprinted from Noda et al. [49]. Copyright 1999, with permission from Elsevier.

spectra. Examples of synchronous and asynchronous contour maps are presented in Fig. 5.4.

It can be recognized from Fig. 5.4 that diagonal peaks are absent in the asynchronous spectrum.

The asynchronous contour map was described as helpful for enhancing resolution in the same way 2D correlation spectroscopy does [12]: Assuming that the cause of the variations of two coincidentally overlapping peaks are sufficiently different, asynchronous correlation peaks will be obtained at the corresponding spectral positions, thus spreading the spectral crowding into two dimensions.

In an NMR experiment, the perturbation dimension usually comprises the nuclear spin frame instead of a time frame or sample frame of macroscopic phenomena such as a chemical reaction. Different phases thus depend on the individual nuclear spins rather than on the molecular assembly of spins. Therefore, the exploitation of phase information within the covariance processing described above might seem little attractive.

### 2.6.2 Applications of Asynchronous Spectra

As an application of using synchronous and asynchronous spectra, Eads and Noda [50] described the covariance analysis of diffusion ordered

spectroscopy (DOSY) NMR spectra of a mixture of compounds. Instead of calculating the DOSY spectrum from a DOSY data set, which would have been a representation of  $^1\text{H}$  NMR spectra against the diffusion rate as indirect dimension, a 2-dimensional correlation plot  $\mathbf{F}(\omega_1, \omega_2)$  was obtained. The diffusion rate information was contained in the signs and intensities of the cross-peaks. The synchronous spectrum displayed cross-peaks between all signals without any discrimination. The asynchronous spectrum showed cross-peaks between signals from molecules whose diffusion rates were different. For interpretation, the comparison of synchronous and asynchronous contour maps was suggested. Analysis of the sign of the cross-peaks was proposed for ranking the molecules according to their diffusion rates.

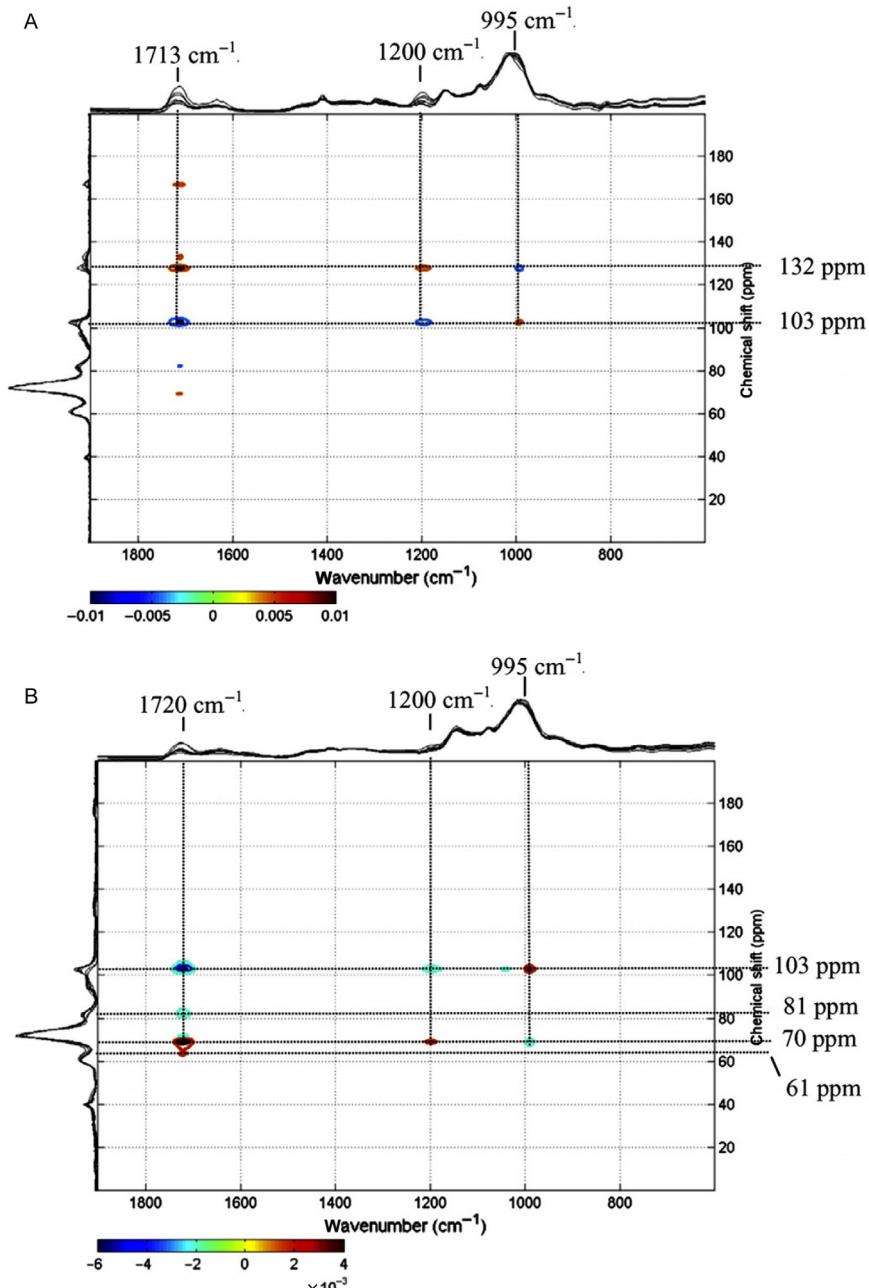
Hu *et al.* [39] used both synchronous and asynchronous maps to investigate silk spider thread by solid-state NMR techniques. They constructed  $^{13}\text{C}$ - $^{13}\text{C}$  correlation maps from a  $^{13}\text{C}$ - $^{13}\text{C}$  spin diffusion experiment that was not further specified. They demonstrated that very few points sampled in the indirect dimensions proved sufficient for their analysis purposes, as the necessary resolution was re-introduced from the asynchronous covariance map.

A consequence of Eq. (5.14) is the possibility of combining heterogeneous spectra as was reported for indirect, unsymmetrical indirect and GIC spectroscopy. Yet, the combination of spectra from different spectroscopic methods is also implicitly contained. Several variants have been summarized [3,48]. As to NMR spectroscopy, Garcia *et al.* investigated the structure of dextrin hydrogels by FTIR and solid-state NMR spectroscopy [51] affording solid-state NMR—FTIR correlation spectra. Changes in sample composition served as the common or perturbation dimension. A synchronous solid-state NMR—FTIR contour map of polymerized dextrin presented in Fig. 5.5.

## 2.7. Heterospectroscopy

Crockford *et al.* utilized the heterospectroscopic methodology to investigate metabolic mixtures [52,53]. They concatenated spectral data from NMR and ultra high-performance liquid chromatography (UPLC) coupled to MS to obtain NMR—MS covariance maps. The approach was baptized SHY as an acronym of statistical heterospectroscopy.

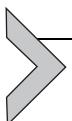
NMR spectra were recorded as 1D  $^1\text{H}$  NMR experiments using a NOESY water suppression sequence with experimental parameters optimized for urine samples in the field of metabolomics. Mass spectra were acquired by



**Figure 5.5** 2D FTIR- $^{13}\text{C}$  NMR correlation spectroscopy: synchronous maps of (A) unpolymerized dextrin and (B) polymerized dextrin. From Garcia et al. [51]. Copyright 2008, with permission from Elsevier.

use of an UPLC-MS system. Two-dimensional histograms were produced thereof with retention time and mass-over-charge,  $m/z$ , ratio as axes. Prior to covariance processing, the UPLC-MS histograms were summed over the whole retention time of a sample such that a pseudo direct injection spectrum was obtained for each sample. The perturbation dimension was represented by the samples subjected to NMR and UPLC-MS analysis. The NMR-MS map correlated the NMR chemical shift with the  $m/z$  ratio, thus connecting NMR and MS information for each molecule. Due to the introduction of a second dimension, a gain in resolution was achieved. An example of a NMR-MS correlation plot is presented in Fig. 5.6.

While the use of direct, indirect and unsymmetrical and generalized covariance is well represented in the literature, the exploitation of heterospectroscopy with NMR as one element is still scarce. Still, covariance analysis as dealt with here is a processing scheme that is open to application. It cannot be excluded therefore that some of the methods are employed without disclosure to a general audience. In this respect, the mathematical or data conversion tools for covariance treatment are readily available today.

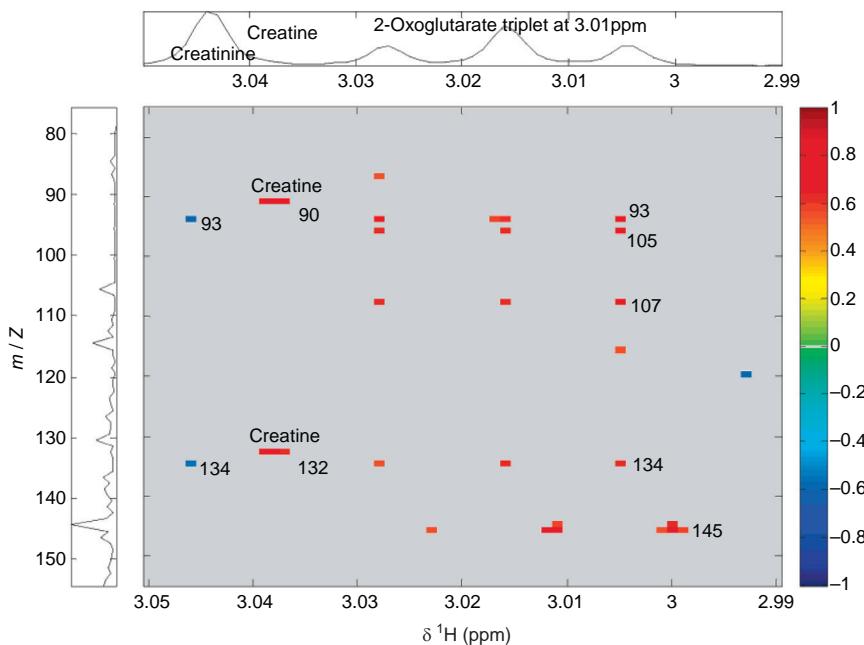


### 3. SOFTWARE FOR COVARIANCE NMR PROCESSING

Reviewing software seems a difficult task since developments often proceed very rapidly, the more when they accompany an area of very active research.

Covariance processing has been implemented into spectrometer manufacturers' software, into spectral analysis and manipulation software. But also research- and problem-oriented solutions have been built. While commercially available solutions are integrated into spectral software packages like Spectrus by ACD/Labs or MNova by Mestrelab Research, open structures are often based on mathematical platforms and interfaces such as MATLAB and provided as freeware or Web-based services. The most accessible software applications for covariance treatment are collected in Table 5.2.

The two large-scale commercial vendors of spectrometer-independent NMR software currently offer covariance processing as a feature of their spectral manipulation and interpretation suites. The Spectrus platform by ACD/Labs and the MNova by Mestrelab Research provide stand-alone solutions that allow the import of all spectrometer data formats without the necessity of pre-processing. Covariance processing can be performed according to the direct, indirect, and generalized indirect formalism. The doubly indirect covariance transformation is found in MNova. While



**Figure 5.6** NMR–MS correlation for control samples expanded to show a 2-oxoglutarate region of the  $^1\text{H}$  NMR spectra (cutoff, 0.55). The correlation of the creatine ions with the creatine NMR singlet is clear. There is a negative correlation of unidentified ions at  $m/z$  93 and 134 to the overlapped creatinine peak, probably deriving from the same compound. The ions correlating directly with the 2-oxoglutarate NMR triplet were not identified but may be related. The ion at  $m/z$  145 correlates with a triplet overlapping 2-oxoglutarate, and it was proposed that it originated from *N*-acetyllysine. Reprinted with permission from Crockford et al. [52]. Copyright 2006 American Chemical Society.

performing the processing is an easy task with both programs and can be automated to some extend, the user cannot implement new features. To the user who is interested in developing covariance variants on his own or to analyse the results of intermediate calculation steps more deeply, it is inevitable to turn to open source solutions if not creating a new program. The toolboxes provided by Snyder and Lafon [38,57] are based on MATLAB by MathWorks, which require a running MATLAB environment. They also need pre-processing of the data by NMRPipe. However, the former toolbox provides all types of covariance processing developed by the Brüschweiler group [57]. Therefore, 4D covariance, Z-matrix formalism including regularization and residual water signal removal can be performed.

**Table 5.2** Collection of software to use for covariance NMR

Software	By (vendor or programmer)	Features with respect to covariance applications	Reference or link
Spectrus	ACD/Labs	Data format independent direct, indirect, unsymmetrical and generalized indirect covariance processing	<a href="http://www.acdlabs.com/products/spectrus/">http://www.acdlabs.com/products/spectrus/</a>
MNova	Mestrelab Research	Data format independent direct, indirect, generalized indirect covariance processing	2014Cob01 [54] <a href="http://mestrelab.com/software/mnova/nmr/">http://mestrelab.com/software/mnova/nmr/</a>
Topspin 3.0	Bruker Biospin AG	Bruker data format; User program (au) for direct and indirect covariance processing	<a href="http://www.bruker.com">http://www.bruker.com</a>
COLMAR	Rafael Brüschweiler, Florida State University and National High Magnetic Field Laboratory	Metabolomics Web portal; NMRPipe, Varian, Bruker data format; direct and indirect covariance processing with optional water removal; also available: DemixC, query, TOCCATA query	[55,56] <a href="http://spinportal.magnet.fsu.edu">http://spinportal.magnet.fsu.edu</a>
2DShige	Shigeaki Morita, Kwansei-Gakuin University	CSV data file; calculation of synchronous and asynchronous maps	<a href="https://sites.google.com/site/shigemorita/home/2dshige">https://sites.google.com/site/shigemorita/home/2dshige</a>
Covariance Toolbox	David Snyder	MATLAB/GNU OCTAVE base; NMRPipe data format; direct, indirect, 4D, generalized indirect and Z-matrix processing	[57] <a href="http://www.mathworks.com/matlabcentral/fileexchange/27264-covariance-nmr-toolbox">http://www.mathworks.com/matlabcentral/fileexchange/27264-covariance-nmr-toolbox</a>

**Table 5.2** Collection of software to use for covariance NMR—cont'd

Software	By (vendor or programmer)	Features with respect to covariance applications	Reference or link
Toolbox	Lafon	[38] MATLAB processing algorithm available upon request	
Nmrglue		Data-independent format; various processing feature; can be used for data preparation for use with Covariance Toolbox	[58] <a href="http://nmrglue.com">http://nmrglue.com</a>
NMRPipe	Frank Delaglio, National Institute of Health	Data independent format; numerous NMR relevant and useful features; can be used for data preparation for use with Covariance Toolbox	[59] <a href="http://spin.niddk.nih.gov/NMRPipe/">http://spin.niddk.nih.gov/NMRPipe/</a>
Other than NMR		FTIR, Raman, UV/vis, etc.	cf. references cited in Ref. [48]

Vendors or distributors and, where available, references are given.

The well-known NMRPipe [59] may be used for data preparation prior to covariance processing by open source solutions. Very recently, Nmrglue was described [58] as another open source solution based on the Python language. Besides several other features, this program may be used as an alternative to prepare data to use with the toolbox. Based on the NumPy library being part of the Python package, covariance calculations were reported to be programmable as well within Nmrglue. As an intermediate solution, the Web server application provided by Brüschweiler and coworkers [55,56] can be understood. Data sets of various formats, cf. Table 5.2, can be uploaded onto a Web server. The user is to select from direct and indirect covariance processing. Since the Web application is part of a suite, further analysis of the data may follow. The server suite was baptized COLMAR and offers spectral fingerprinting of individual compounds if the data were derived from chemical mixtures. The fingerprint spectra may subsequently

be screened against a NMR database for identification. The parts of the suite were termed DemixC, query and TOCCATA query [56,60,61]. They were intended for the deconvolution of metabolic mixtures and the identification of the individual compounds.

Among the programs not particularly dedicated to use in the field of NMR spectroscopy, we refer to 2DShige. The program was devised by S. Morita and may be accessed for download via <https://sites.google.com/site/shigemorita/home/2dshige>. The covariance transformations contained therein are based on the work by Noda. The program computes the synchronous and asynchronous maps from data in CSV format. This program may be useful for heterospectroscopy applications. For an overview of software available in the fields of spectroscopy not related to NMR, the reader is referred to a review by Noda and references cited therein [62].

In conclusion, both commercial and non-commercial software solutions are available for the processing of NMR spectra or mixed time–frequency matrices with covariance methods. The advantage of commercial programs is their ease of use concerning installation and data preparation, for example, data obtained from spectrometer software can directly be imported, whereas non-commercial packages may require a platform such as MATLAB and a data preparation utility, for example, NMR Pipe, to be available and running on the processing computer. The provision of Web-based services renders the processing user-friendlier and removes the requirement of pre-installed mathematical platforms. For-profit organizations may yet be reluctant to upload their data to servers beyond their firewalls, especially data from biological or metabolic studies.

The advantage of the open source solutions is their greater transparency towards the applied mathematical operation and the result of each step and their greater flexibility to develop or implement novel or own covariance processing schemes at need. A comparison of the results of the various software programs, that currently allow covariance processing on a defined selection of spectra or data sets, would be highly desirable.



#### 4. COVARIANCE AND NUS—THE COMBINATION OF TWO APPROACHES TO FAST METHODS

An extremely intriguing and—in the authors' opinion—promising aspect of covariance NMR lies in the combination of non-uniform sampling (NUS) schemes and covariance processing. Both techniques strive for the reduction of experimental times or an increase in resolution and/or

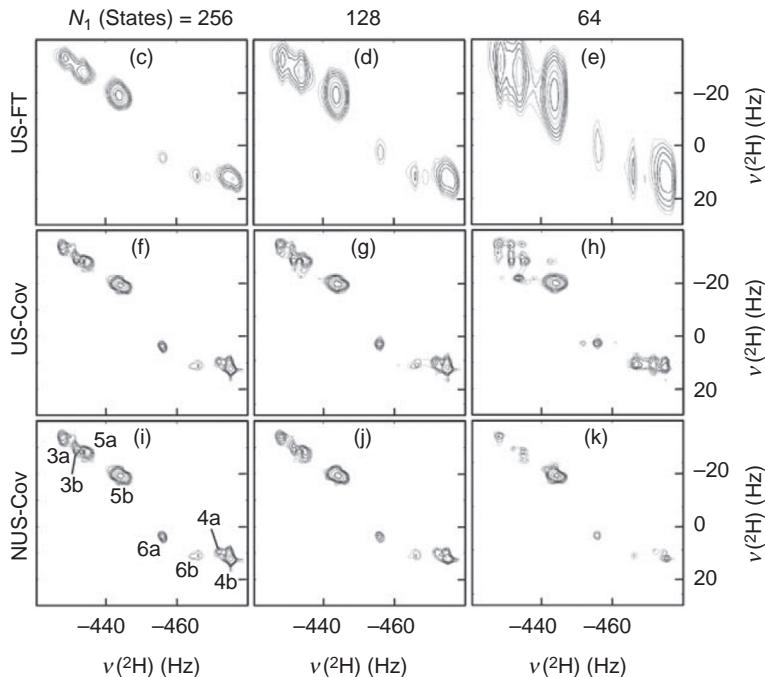
sensitivity. Since it exceeds the scope of this chapter to review NUS, the interested reader is referred to the recent overview by Hyberts *et al.* [63]. Applications of sampling the indirect dimension of a 2-dimensional NMR experiment other than in a linear manner were described in the early 1990s [64]. The research on NUS schemes experienced intensification during the last decade [65–67]. The leading principle of NUS is the reduction of data acquired in indirect dimensions. While in ordinary 2D NMR spectra the indirect domain is divided into increments spaced by the dwell time, a NUS scheme may use exponentially, radially or randomly distributed increments. As the Fourier transform requires equidistant data points, the omitted data points need to be left zero, or the values need to be calculated, that is, emulated or reconstructed. Although Hu *et al.* [39] pointed out the advantages of using direct covariance processing on sparsely sampled data, no indices are found that they used NUS schemes at that time. Chen *et al.* [68] demonstrated that direct covariance processing in principle is not limited by the Nyquist theorem with respect to incrementing the indirect dimensions. Yet, undersampling leads to the appearance of parallel diagonals in covariance spectra. A diagonal on the side of the main diagonal is apparent and duplicated ( $n - 1$ )-fold for  $n$ -fold undersampling. Since the relative frequencies of the diagonals can be calculated from spectral parameters, the diagonals can be removed by appropriate masking or filtering. It was observed that undersampling would be suitable for COSY experiments whereas a minimal sampling scheme obeying the Nyquist theorem would give better results for TOCSY spectra [68]. As a first step towards NUS and covariance processing, Chen *et al.* selected a series of FIDs from a larger data set to construct their limited test data set. To this purpose, they started with the first time increment for the first sparse data set and with the 25th increment for a second data set. The resulting data sets were successfully converted into covariance maps having the high resolution of the direct dimension. Expectedly, no first-order phasing problems were observed in neither case.

After the proof-of-principle that covariance proves robust in the case of sparsely sampled data sets, non-uniformly sampled data were finally subjected to covariance processing for the analysis of anisotropic liquid crystal samples investigated by natural abundance 2D  $^2\text{H}$  NMR [38]. A natural abundance deuterium (NAD) Q-COSY with a z-gradient filter, Q-COSY Fz, was used for data acquisition [69,70]. The sequence involves cancellation of diagonal peaks. As NUS scheme, exponentially weighted deterministic on-grid sampling was employed [71]. Yet, quantitative direct

covariance processing demands the presence of cross-peaks [17]. Thus, Lafon *et al.* [38] suggested a modified regularization method, since the addition of a unity matrix multiplied by a prefactor as described previously by Chen *et al.* [17], requires 2D Fourier transformation when applied to NUS data. This altered method introduces the diagonal peaks at the stage of the mixed time–frequency data matrix. Hence, the lack of diagonal signals can be compensated for by the modified regularization procedure, allowing quantitative covariance processing on non-uniformly sampled sparse data. As an example for probing various combinations of uniform and NUS together with Fourier transform and covariance processing approaches, a selected region of the resulting contour plots of the NAD Q-COSY Fz spectra (Fig. 5.7) is presented [38].

As a result of that study, the combination of covariance processing and NUS afforded a decrease in experimental time by a factor of two relative to linear sampling and covariance processing. It even shortened the total acquisition time by a factor of four as compared to conventionally sampled data and FT processing. Again, enhanced resolution and improved sensitivity were observed. The proof of concept was predicted to be amenable to other homonuclear 2D NMR experiments on liquid, aligned or solid samples.

Another study on covariance NMR associated with NUS was reported in 2012. Li *et al.* [72] combined 2-dimensional covariance spectroscopy (COV2D) with a non-uniform and consecutive acquisition scheme (NUCA). Through this acquisition scheme, a decreasing number of accumulations were recorded as a function of  $t_1$ , that is, the increments in the indirect dimension of a 2D experiment [72–75]. The increments were sampled consecutively and equidistantly. The NUS scheme was implemented into a  $^{13}\text{C}$  DARR experiment with L-[U- $^{13}\text{C}$ ]-histidine hydrochloride and the proteins [U- $^{13}\text{C}$ ]-GB1 and [U- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-protorhodopsin. Within the study, the NUCA-COV2D method performed well for a number of  $t_1$  points reduced by a factor of 1.5–3 without any line broadening as compared to FT 2D processing. The total experiment time could be shortened by a factor of 3–6 due to an improved signal-to-noise ratio. Further analyses included the indiscriminate decrease of  $t_1$  increments, which eventually diminished the spectral resolution and the signal-to-noise ratio. An optimum of  $t_1$  points was observed, which depended on the minimum line width in the FT 1D spectrum and the dwell time of the indirect dimension. In order to ensure acceptable spectral parameters it was suggested to determine a safety margin or threshold of  $t_1$  points for each experiment.



**Figure 5.7** Natural abundance deuterium (NAD)  $^2\text{H}\{^1\text{H}\}$  Q-COSY-Fz FT spectrum of methyl vernoleate dissolved in a mixture of the liquid crystal PBLG and  $\text{CHCl}_3$  at 305 K. The assignment of  $^2\text{H}$  quadrupolar doublets is indicated in the bottom left panel. The NAD signal of  $\text{CHCl}_3$  is not shown. The same region obtained by different sampling and processing methods is shown in the squares. Top row: uniform sampling and FT. Middle row: uniform sampling and covariance. Bottom row: NUS and covariance. Left column: spectra were processed using 256  $t_1$  increments. Middle column spectra were processed with 128  $t_1$  increments. Right column: spectra were processed with 64  $t_1$  increments. For the four panels on the upper right, the selected data points corresponded to the shortest  $t_1$  times. The two panels to the right of the bottom row, the data points followed an exponentially decaying distribution. Reprinted with permission from Lafon et al. [38]. Copyright 2011 Wiley-VCH Verlag GmbH & Co., KGaA Weinheim.

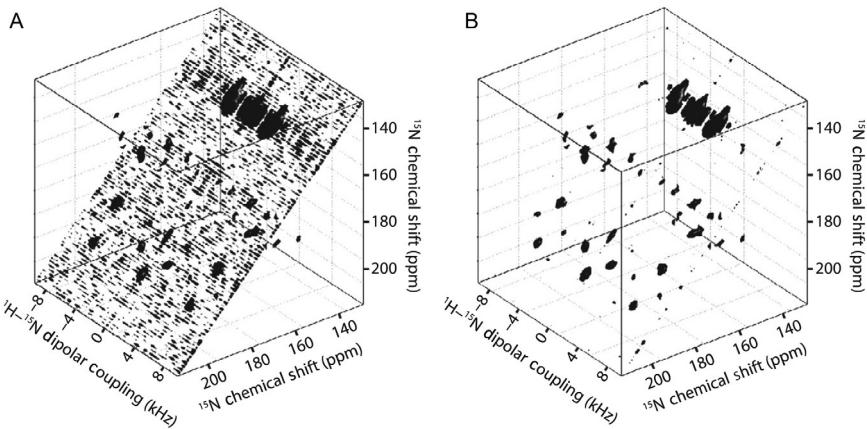
A covariance study by Li *et al.* [76] dealt with the comparison of various sampling schemes and accumulation profiles with respect to their usability with covariance processing. The authors divided sparse sampling schemes into the non-uniform and the  $t_1$  cut-off (CUO) schemes. While common NUS schemes employ—exponentially—increasing  $\Delta t_1$  spacing as a function of  $t_1$ , a cut-off design implies the sole acquisition of signals at short  $t_1$  times up to a maximum value. The 2D NMR spectrum is then reconstructed with appropriate methods. Various standard, NUS and

DUO schemes were tested together with a SHANGHAI  $^{13}\text{C}$ - $^{13}\text{C}$  recoupling sequence on [ $\text{U}-^{13}\text{C}$ ] histidine hydrochloride and together with a  $^{13}\text{C}$  2D DARR experiment on [ $\text{U}-^{13}\text{C}$ ]-GB1 protein and [ $\text{U}-^{13}\text{C}, ^{15}\text{N}$ ]-proteorhodopsin, cf. previous study [72]. In their study, Li *et al.* found that CUO-COV2D proved more sensitive than NUS-COV2D for a given experimental time and homonuclear correlation spectroscopy represented by the two solid-state NMR experiments used. That finding suggested that the signal corresponding to short evolution times was the most sensitive within covariance processing and that no resolution was lost. The cut-off strategy could also be combined with a Gaussian accumulation profile as in NUCA, cf. above. The signal-to-noise ratio could thus be improved further. Systematic variation of CUO and NUCA was also investigated. A maximum reduction of  $t_1$  increments by a factor of 6–12 was achieved. However, the validity of the findings was limited to exponentially decaying 2D signals, thus excluding sine/cosine modulated or constant-time signals.

Interestingly, NUS and covariance seems up-to-now to be primarily attractive to solid-state NMR methodology. Along these lines, Lin and Opella [77] extended the range of application towards multidimensional NMR experiments. The 3D homonuclear spin-exchange/separated-local-field (SLF) spectrum of the  $^{15}\text{N}$ -labelled Pf1 coat protein in magnetically aligned macrodiscs [78] served as an example. By this method, the spectral reconstruction from compressed sensing (CS) [79] and covariance processing proved successful with only 11.7% sampling without loss of resolution as compared to 100% sampling of traditional sampling. The reconstructed  $^1\text{H}$ - $^{15}\text{N}$  SLF- $^{15}\text{N}$ / $^{15}\text{N}$  spin-exchange 3D spectrum of a single crystal sample of  $^{15}\text{N}$  N-acetyl leucine is given in Fig. 5.8.

The  $^1\text{H}$ - $^{15}\text{N}$  dipolar coupling dimension and the  $^{15}\text{N}/^{15}\text{N}$  correlations were reconstructed by CS and covariance processing, respectively. Prior to CS reconstruction, optimized NUS schemes were applied. The diagonal noise was minimized through subtraction of average values during the covariance processing. The combination of both experiments allowed to obtain correlations, sizes of dipolar couplings and chemical shifts from a single spectrum. With respect to the benefits of high-resolution and intrinsically high sensitivity, the potential of the method for use in solid-state NMR structure determination of biopolymers was emphasized.

Although the combination of NUS and covariance processing has only been reported for solid-state NMR analyses so far, we expect its use and exploitation for solution-state NMR to start soon. Especially in laboratories



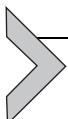
**Figure 5.8** (A) The reconstructed  $^1\text{H}$ - $^{15}\text{N}$  SLF  $^{15}\text{N}/^{15}\text{N}$  spin-exchange 3-dimensional spectrum of a  $^{15}\text{N}$ -labelled  $^{15}\text{N}$   $\text{N}$ -acetyl leucine single crystal. (B) The same spectrum without diagonal noise, which was removed by subtracting their average value. Covariance processing was applied to  $^{15}\text{N}/^{15}\text{N}$  correlations for the reconstruction. Spectra were recorded with 1024 complex data points in the direct dimension. 50 complex points were acquired in the indirect dimension using the alternate phase sampling scheme, which corresponds to 23.4% sampling, and the dipolar coupling dimension was reconstructed from 50% sampling (44 real points) by compressed sensing. *From Lin and Opella [77]. Copyright 2014, with permission from Elsevier.*

operating under time constraints, in particular commercial organizations, where the reduction of experimental time of 2D experiments is of vital interest, NUS acquisition and covariance processing should be investigated. In this context, the availability of robust and easy to use software to process NUS data with covariance methods with a high degree of automation will be a prerequisite. For routine laboratories, a successfully automated workflow may provide a generic 2D or 3D acquisition and processing example or method containing a default NUS sampling scheme and the appropriate reconstruction and covariance formalism.

It can be argued that within the distinction between the more methodologically oriented first part and the following applied part of this chapter the occurrence of NUS in combination with covariance processing in the first part is somewhat arbitrary. The authors acknowledge in this respect that from the topics to follow, mixture deconvolution, specific long-range  $J$ -correlation covariance maps, and pure-shift covariance NMR may prove evenly important and are at present even wider applied than NUS

covariance NMR. Yet, NUS participates directly in the quest for more sensitivity or higher resolution, and was therefore placed before [Section 5](#).

The applications of covariance NMR to chemical compounds and their mixtures described in the scientific literature will be distinguished and reviewed according to the four types of transformations that are specific for NMR. An additional cluster is given for the general covariance treatment with reference to the synchronous and asynchronous maps.



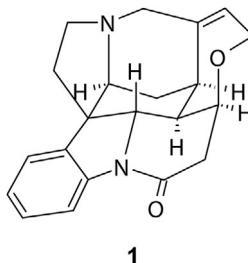
## 5. APPLICATIONS OF COVARIANCE NMR TO SMALL MOLECULES

### 5.1. Direct Covariance

The application of direct covariance processing to either NMR spectra or mixed time–frequency data matrices was preferentially used for spectral enhancement with respect to resolution. As the resolution of the indirect or incremented dimension is endowed with that of the direct or acquisition dimension, homonuclear spectra were treated almost exclusively.

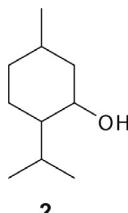
#### 5.1.1 Solution State

In their introductory description and study of direct covariance NMR, Brüschweiler and Zhang [\[37\]](#) investigated a sample of ubiquitin at 600 MHz with H–H TOCSY and H–H NOESY experiments. After covariance treatment of the  $2048 \times 1024$  point data matrices, they found in agreement with theory that indeed high resolution was obtained in the indirect dimension. No separate phase correction in that dimension was required, either. Furthermore, it was proven that covariance transformation proved robust against incompleteness of FIDs or removal of single FIDs, for example, in cases of receiver gain overflow or data corruption. The covariance matrix is by itself symmetric, thus symmetrization procedures did not need to be performed. Ubiquitin was further studied by 2QF-COSY, yielding comparable results [\[35\]](#). Covariance computation not only rendered phase correction in the indirect dimension unnecessary but apodization as well. On analyzing H–H COSY spectra of strychnine **1**, it was recognized that the covariance map afforded multistep correlations, further termed RCOSY [\[45\]](#). These correlations were interpreted in terms of relayed coherences. For example, COSY signals originating from magnetization transfers A to B and B to C lead to direct correlation signals A to C in the covariance COSY spectrum.

**1**

It was suggested that these long-range correlations might be used to validate weak responses in H–H TOCSY spectra. The spectral data used prior to covariance treatment were of the array size  $2048 \times 128$  points and were recorded within 30 min. In order to ensure the presence of peak volumes in accordance to 2D FT H–H TOCSY spectra, the regularization procedure was devised by Chen *et al.* [17]. As model compounds, strychnine and the cyclic decapeptide antamanide were analysed at 800 MHz. The TOCSY mixing time was varied to test the regularization procedure against different amounts of magnetization transfer. Depending on the strength of the regularization, star-like peak shapes were observed. As a remedy to this problem, the regularization scaling parameter could be optimized for different regions of the spectrum.

A study was conducted which did not employ a frequency symmetric experiment prior to direct covariance processing. Zhang *et al.* [80] subjected a 2D C–C INADEQUATE data set of  $2048 \times 1024$  data points to covariance NMR. They thus transformed the common double-quantum representation of an INADEQUATE spectrum into a symmetric single-quantum map. The resulting symmetric spectrum was considered easier to interpret. Menthol **2** was used as reference compound.

**2**

Further to the spectral improvement, covariance NMR was utilized as a tool to identify individual compounds in mixtures. The recognition of highly selective patterns stored as columns or rows of covariance maps, their extraction followed by a database search and the assignment to an individual

compound were assembled into an automated or semi-automated program, that was baptized DemixC [21,81,82]. The program first started out as a procedure to deconvolve an NMR spectrum of mixtures without prior physical separation [21]. The H–H TOCSY matrix of  $1024 \times 1024$  data points was understood as a series of 1D TOCSY spectra recorded on a mixture of glycine, leucine, lysine and valine. These spectra could be extracted as rows of the matrix and screened against a database in a semi-automated mode. The applicability in the fields of metabolomics was suggested. The next report on DemixC and direct covariance TOCSY included two mixtures of glycine, lysine, valine and glycine, leucine, lysine, valine at 600 and 800 MHz [83]. In addition, antamanide was studied as it represents a peptidic example with significant resonance overlap due to its amino acid composition. A similarity as elements of the matrix  $\mathbf{O}$  between each row vector and column vector of the covariance spectrum  $\mathbf{C}$  was determined according to Eq. (5.32) associated with an importance index  $P$  for the amount of overlap.

$$\mathbf{O} = \mathbf{C}^T \cdot \mathbf{C} \quad (5.32)$$

Within the importance vector, overlap is due to spins of the same spin system. The importance vector resembles signals of a 1D NMR spectrum with the signal intensities the higher the more similar the corresponding rows. The approach proved robust and also provided the information of 1D NMR. The method was regarded amenable to semi-automated side-chain assignment for peptides or small proteins.

Covariance treatment of TOCSY spectra with subsequent submission to DemixC differs from statistical total correlation spectroscopy (STOCSY) since it is based on spectral variables,  $t_1$ , instead of different samples, that is Cov<sub>vv</sub> versus Cov<sub>ss</sub> in the nomenclature of Noda, where Cov denotes the covariance generated map, v the variable and s the sample dimension.

In a subsequent study, the DemixC method was applied to a mixture of D-glucose, L-histidine, L-lysine, serotonin hydrochloride, D-sorbitol as well as to the venom of the walking stick insect *Anisomorpha buprestoides*, which consists of at least six compounds [84]. The H–H TOCSY spectra of  $2048 \times 512$  points were recorded in a screening setup at 600 MHz with a 1 mm probehead. With the help of databases, the mixture was deconvolved and the venom identified. Thus, automated mixture deconvolution in screening mode using covariance processed H–H TOCSY experiments was found feasible.

While the TOCSY experiments provide almost a 1D  $^1\text{H}$  NMR spectrum along its rows due to long-range magnetization transfer, a H–C HSQC spectrum profits from the resolution of the large  $^{13}\text{C}$  spectral dispersion. Thus, Zhang *et al.* generalized the DemixC to H–C HSQC–TOCSY to take advantage of both experiments [85]. A metabolic test mixture of carnitine, glucose, lysine, myo-inositol and shikimate as well as an extract from a human prostate cancer cell line, DU145, served as study objects. The HSQC–TOCSY spectra were recorded with a resolution of  $2014 \times 1024$  data points at 800 MHz using cryoprobe technology. To take full advantage of the experimental data, both direct and indirect covariance processing was applied before the resulting maps were submitted to DemixC. A lower signal-to-noise ratio for HSQC–TOCSY was reported as a disadvantage. Nevertheless, the narrower line width of the  $^{13}\text{C}$  signals as compared to the  $^1\text{H}$  traces and the additional identification of the compounds by  $^{13}\text{C}$  resonances were counted as advantages.

A very recent report on the application of DemixC will be summarized below, since the spectra were processed via indirect covariance [81].

### 5.1.2 Solid State

As is the case for liquid samples, solid-state NMR experiments profit from resolution enhancement. Yet, the reduction of increments in the indirect dimension leading to shorter experiment times might be of greater interest than the increase in resolution itself due to the intrinsically broad lines. A study was performed on silicate, sodium borosilicate glass, aluminophosphates, and  $^{13}\text{C}$ -labelled tyrosine [32]. Experiments were recorded to yield  $^{29}\text{Si}$ – $^{29}\text{Si}$  INADEQUATE,  $^{31}\text{P}$  RFDR and  $^{13}\text{C}$ -Post-C7-DQ spectra. The gain in experiment time was always reported as a factor of 10 as compared to the acquisition of FT processed data sets, since only an optimized number of  $t_1$  slices needed to be recorded. It was thus possible to study a silicate network of glasses by 2D NMR in a reasonable time frame despite the long relaxation times of  $^{29}\text{Si}$ .

Covariance spectra with 512 and 256 increments in the indirect dimension from  $^{13}\text{C}$ -labelled microcrystalline proteins were visually inspected with respect to their signal-to-noise and resolution enhancement [86]. The C–C correlation spectra with PARIS recoupling were recorded on the  $^{13}\text{C}$ - and  $^{15}\text{N}$ -labelled protein Crh which is an 85 amino acid domain-swapped homodimer of catabolite repression phosphocarrier protein of approximately 21 kDa. The spectra were judged to display no difference in resolution or sensitivity. Thus direct covariance processing was

found suitable and advantageous in solid-state NMR protein investigations. Analogous conclusions were drawn from a binding study on the 42-residue amyloid  $\beta$ -protein A $\beta$ 42, that was selectively labelled, and curcumin [87]. The latter is believed to prevent the pathogenesis of Alzheimer's disease. The A $\beta$ 42-fibrils were exposed to dipolar-assisted rotational resonance (DARR), a 2D distance-related solid-state NMR experiment under magic-angle spinning conditions. Assuming a fixed experiment time, covariance processing led to an improved signal-to-noise ratio, since the reduction in  $t_1$ -increments could be invested in an increase in accumulations. A cross-peak dependence on the mixing time was found, where the signal-to-noise of a cross-peak increased with prolonged mixing time.

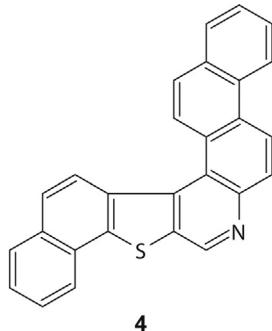
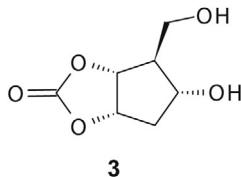
## 5.2. Indirect Covariance

When regarding the matrix notation of the covariance calculations, it is evident that not only the direct dimension of a spectrum but also the indirect dimension can be turned into a symmetric correlation plot, cf. Eqs. (5.15) and (5.16). Due to current probe technology and pulse sequence design, the direct dimension in solution is mostly represented by the proton frequency, whereas the indirect dimension can arise from a variety of nuclei. In contrast, solid-state NMR experiments often rely on the detection of nuclei other than protons due to a preferable line width-to-dispersion ratio. The following section summarizes the covariance-based examples so far.

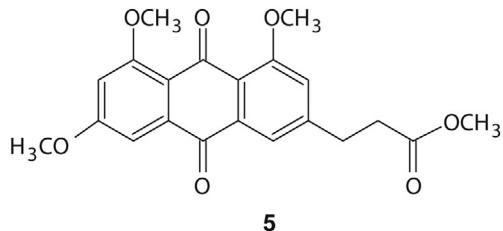
### 5.2.1 Solution State

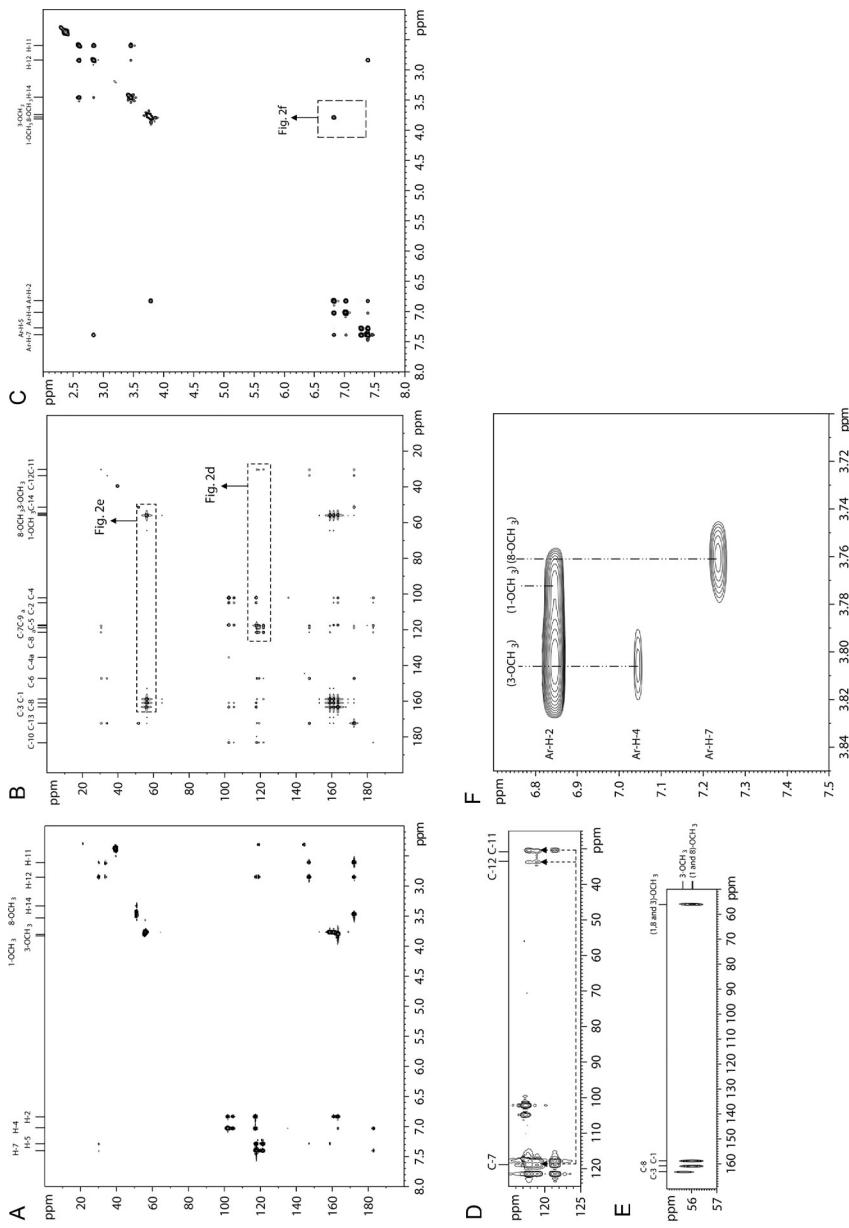
The first transformation of a 2D NMR experiment into an indirect contour plot was reported by Zhang and Brüschweiler [21]. They acquired one H–C HSQC–TOCSY data set of a mixture of glycine, lysine, valine and another one of the  $^{13}\text{C}$ -labelled cyclic peptide antamanide at 600 MHz. They computed the equivalent of a C–C TOCSY or C–C correlation map. Enhanced sensitivity as compared to a hypothetically direct detected C–C TOCSY was quickly revealed and estimated to be 8:1 according to the gyromagnetic ratio of proton and carbon nuclei. Inspired by that study, Blinov *et al.* [18] investigated in depth the occurrence and recognition of artefacts in C–C correlation spectra representations generated from conventional H–C HSQC–TOCSY and inverse direct response (IDR) HSQC–TOCSY experiments. The C–C COSY or TOCSY maps of cyclopentafuranone **3** and naphtho[2',1':5,6]-naphtho[2',1':4,5]thieno[2,3-c]quinolone **4** revealed artefacts that were baptized type I and type II artefacts. Compound **3** served as an example for a well-resolved spectrum while compound **4**

displayed a heavily crowded spectrum. The artefact types were traced back to signal overlap in the parent spectrum either from accidental isochrony or from relayed responses. It was shown how a symmetrization procedure in combination with a separate treatment of the phase-separated signals from the IDR HSQC–TOCSY could help to minimize the problem of false correlations.

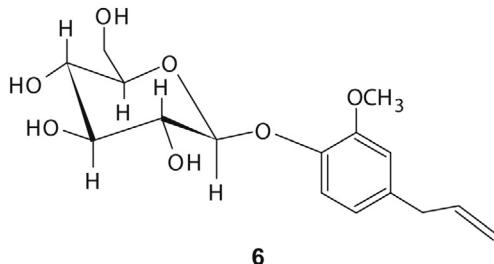


Both direct and indirect covariance processing was applied to H–C HMBC data, although the HMBC was described as a long-range HSQC with  $^1J_{\text{CH}}$  decoupling [88]. As model systems, an emodin derivative **5** and eugenol- $\beta$ -D-glucopyranoside **6** were chosen. Obviously, covariance treatment afforded a H–H and a C–C correlation plot. The spectra are presented together with the parent H–C HMBC in Fig. 5.9.





**Figure 5.9** 2D NMR spectra of the emodin derivative **5**: (A) H–C HSQC spectrum with <sup>1</sup>H-assignments given along the top edge. (B) Indirect C–C HMBC covariance spectrum with <sup>13</sup>C assignments given along the top edge. The dashed boxes indicate the positions of the enlarged regions shown in (D) and (E); (C) Indirect <sup>1</sup>H-HMBC covariance spectrum with <sup>1</sup>H-assignments given on the top edge. The dashed box indicates the location of the enlarged region shown in (F); (D) Expanded region from (B), containing cross-peaks from –OCH<sub>3</sub> to aromatic <sup>13</sup>C (ar-C  $\leftrightarrow$  –OCH<sub>3</sub>); (E) Expanded region from (B), containing cross-peaks between the aromatic C–Cs of the –CH<sub>2</sub>COOCH<sub>3</sub> moiety (CH<sub>3</sub>; 33.5 ppm and CH<sub>2</sub>; 30.6 ppm); (F) Enlarged region from (C) showing cross-peaks in the aromatic/OCH<sub>3</sub> region of the <sup>1</sup>H-HMBC covariance spectrum. Reprinted with permission from Schöfberger et al. [88]. Copyright 2007 John Wiley & Sons, Ltd.

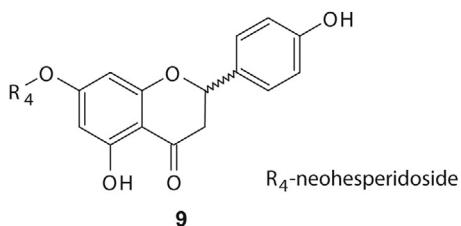
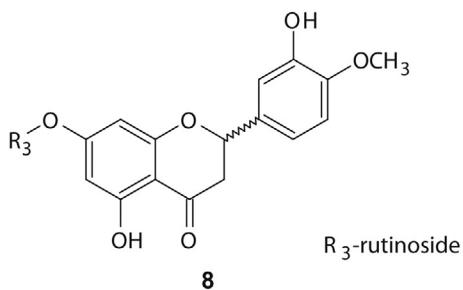
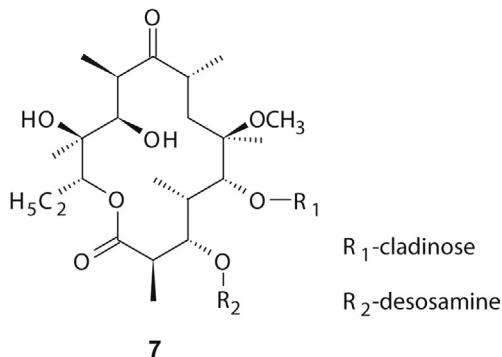


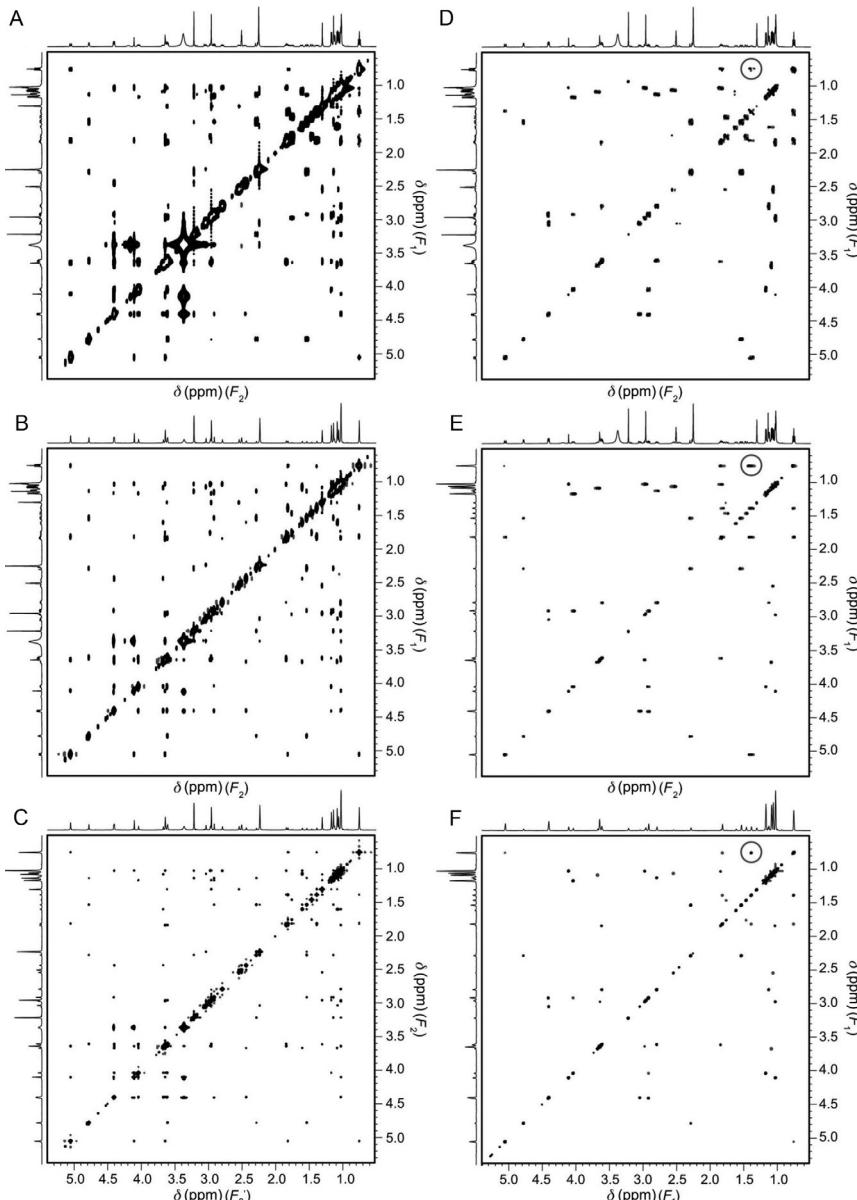
All expected signals were present in the processed spectra. The transfer of the signals to connectivity maps facilitated the interpretation. The intensity of cross-peaks was found dependent on the number of common heteronuclear coupling partners and on the amplitudes of the corresponding cross-peaks in the heteronuclear long-range correlations spectra. As a consequence, nuclei with vanishing mutual couplings could be correlated through a common coupling partner. This was suggested helpful for conformational analyses.

A different goal was aimed at when H–H NOESY and H–H TOCSY data sets were processed by indirect covariance NMR [47]: The transformation was used to remove strong residual water signals from the spectra of a ubiquitin sample in water. The experiments were acquired with  $1024 \times 256$ , TOCSY,  $1024 \times 512$ , NOESY and ROESY, data points at 800 MHz. The data were zero filled to  $2\text{ k} \times 2\text{ k}$  points before covariance processing. The investigators showed two ways to eliminate the water strip, that is, the  $t_1$ -noise of the water signal. The first manner was indirect covariance transformation, the second was to set the data matrix columns to zero before covariance treatment. While both approaches work for TOCSY and NOESY spectra, they fail for ROESY spectra due to the much stronger and broader water signal. For H–H COSY experiments that have dispersive diagonal peaks, the calculations would not be effective either, since the dispersive peak forms have adverse effects on matrix multiplication. Instead, the use of 2QF-COSY experiments was advised. As drawback of the method, the loss in resolution due to the mapping of that of the indirect dimension onto the direct dimension was pointed out.

A desired loss in resolution was the driving force in a recent study on the macrolide antibiotic clarithromycin **7** at 400 MHz [89]. A H–H TOCSY sequence with a Zangerer–Sterk module [90] for suppressing homonuclear  $J$ -evolution in the indirect dimension formed the basis for a doubly pure-shift TOCSY. The Fourier transform in the direct dimension was followed

by indirect covariance such that the coupling patterns present in the direct dimension were collapsed. The same strategy was pursued and extended to NOESY and COSY in a subsequent study [91]. Again clarithromycin, but also a mixture of hesperidin **8** and naringin **9** served as examples. Experiments were performed at 500 MHz with  $16\text{ k} \times 128$  data points for H–H NOESY and  $16\text{ k} \times 1\text{ k}$  for H–H COSY. The data matrices were truncated or zero filled to  $4\text{ k} \times 4\text{ k}$ , NOESY, and  $4\text{ k} \times 2\text{ k}$ , COSY, prior to indirect covariance processing. As noticed above, the covariance calculation led to the removal of the multiplet structure in the former direct dimension. At the same time, undesired anti-phase terms were eliminated. The covariance and the Fourier transformed spectra are presented in Fig. 5.10 to illustrate the superior resolution and simplicity.





**Figure 5.10** NOESY (A–C) and 2QF-COSY (D–F) spectra of clarithromycin **7** in  $\text{DMSO}-d_6$ : (A) conventional NOESY, (B) Zanger-Sterk pure-shift NOESY, (C) covariance spectrum of parent data set 2 (B), (D) conventional 2QF-COSY spectrum, (E) constant-time 2QF-COSY and (F) covariance processing of data set (E). Reprinted with permission from Aguilar et al. [91]. Copyright 2012 Wiley-VCH Verlag GmbH & Co., KGaA Weinheim.

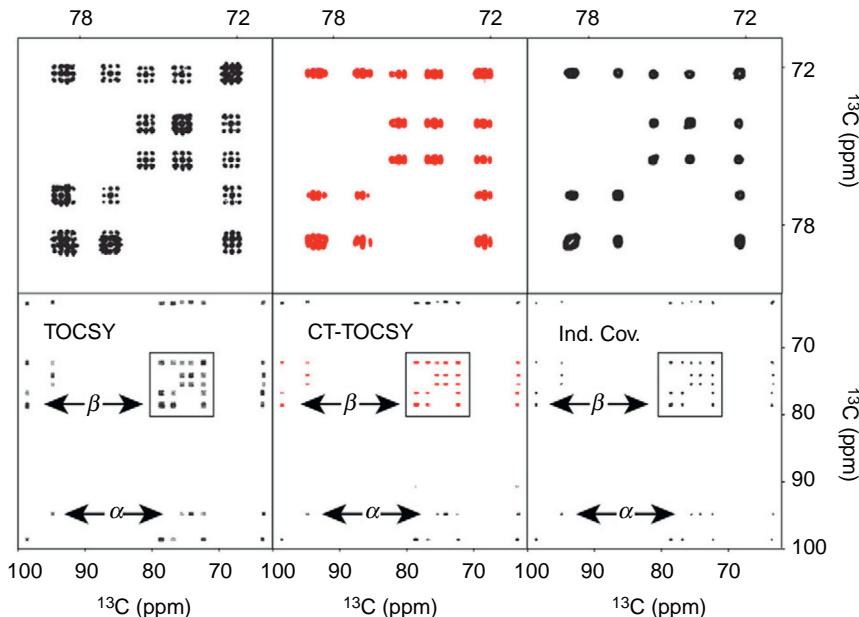
The reduction in signal density and hence overlap made possible the distinction of the flavonoids hesperidin and naringin, that are present in citrus fruits. They occur as diastereoisomers whose proportions depend on the ripeness of the fruit. Aguilar *et al.* predicted that the most common 2D NMR experiments could be adapted to decoupling methods of Zangerer–Sterk of constant-time type such that pure-shift 2D correlation spectra could be produced. They would be of enormous value in structure elucidation and analyses of mixtures.

The same strategy was applied to acquire a C–C TOCSY on  $^{13}\text{C}$ -labelled glucose and a  $^{13}\text{C}$ -labelled algal amino acid mixture and to transform it into a homo-decoupled correlation map [81]. To this purpose, a constant-time TOCSY of  $1024 \times 512$  data points was recorded at 700 MHz using a cryo-probe. The constant-time element in the pulse sequence suppressed the evolution of couplings in the indirect dimension. Prior to covariance processing line broadening functions had to be applied since cross-peaks whose multiplets had no signal at the centre frequency would have disappeared during the symmetrization. The use of a C–C TOCSY circumvents the drawbacks of a H–C HSQC that lacks the correlations of spin systems and is therefore not suitable for molecule identification in mixtures. The resolution enhancement factor as compared to a standard TOCSY was reported as 4, the line width shrinking from 70 to 17 Hz. The spectral improvement from the 2D TOCSY with FT processing, the constant-time TOCSY with FT processing and the constant-time TOCSY with indirect covariance processing according to Zhang *et al.* [81] is demonstrated in Fig. 5.11.

### 5.2.2 Solid State

The generation of indirect homonuclear correlation spectra was described by Hu *et al.* [92] for a  $^{27}\text{Al}$ – $^{31}\text{P}$  SPAM MQ-J-HETCOR experiment, which was transformed into a *J*-coupled Al–Al correlation spectrum or the equivalent to a 2D HOMCOR H–HSQC. The sample studied consisted of aluminium phosphate in polyhedral mode,  $\text{AlPO}_4\text{-}14$ . Due to the symmetric correlation representation, homonuclear connectivities were regarded easier to interpret.

A more recent study employed both direct and indirect covariance on  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled histidine in a polycrystalline sample [34]. The investigators constructed C–C and N–N correlation maps from a H–H dipolar mediated  $^{13}\text{C}$ – $^{15}\text{N}$  shift correlation experiment where the  $^{13}\text{C}$  and  $^{15}\text{N}$  FIDs were simultaneously acquired with multiple receivers. The correlation plots obtained were equivalent to CHHC, NHHN and NHHC spectra. Significant



**Figure 5.11** 2D C–C TOCSY (left), C–C CT-TOCSY (middle) and indirect covariance C–C constant-time TOCSY spectra (right) of  $^{13}\text{C}$ -labelled glucose. The three top panels depict expansions of the boxed spectral regions in the lower panels. The double arrows indicate selected cross-sections that belong to the  $\alpha$  and  $\beta$  forms of glucose. Reprinted from Zhang et al. [81]. Copyright 2012, with permission from Elsevier.

fewer data points were required to obtain a well-resolved spectrum from covariance processing in comparison to conventional Fourier transformation.

From the indirect covariance of a single spectrum, the overview moves towards a much more intense investigated and applied field: unsymmetrical and GIC NMR. Since the introduction of transforming two heteronuclear spectra into the spectrum of a hyphenated experiment by UIC or UIDC [19,22], the added value of this visualization scheme was quickly recognized and many spectral variants were created. When the formalism was eventually described that allowed the square root operation to a UIC correlation spectrum, which was baptized general indirect covariance [14], the advantages and the closer formal similarity to FT spectra seem to have continuously shifted the focus from UIC towards GIC. The qualitative results of both transformations are nevertheless rather comparable. The following section will guide through the applications of UIC and GIC.

The application of doubly indirect covariance was only described for a study on a metabolic model mixture consisting of isoleucine, carnitine,

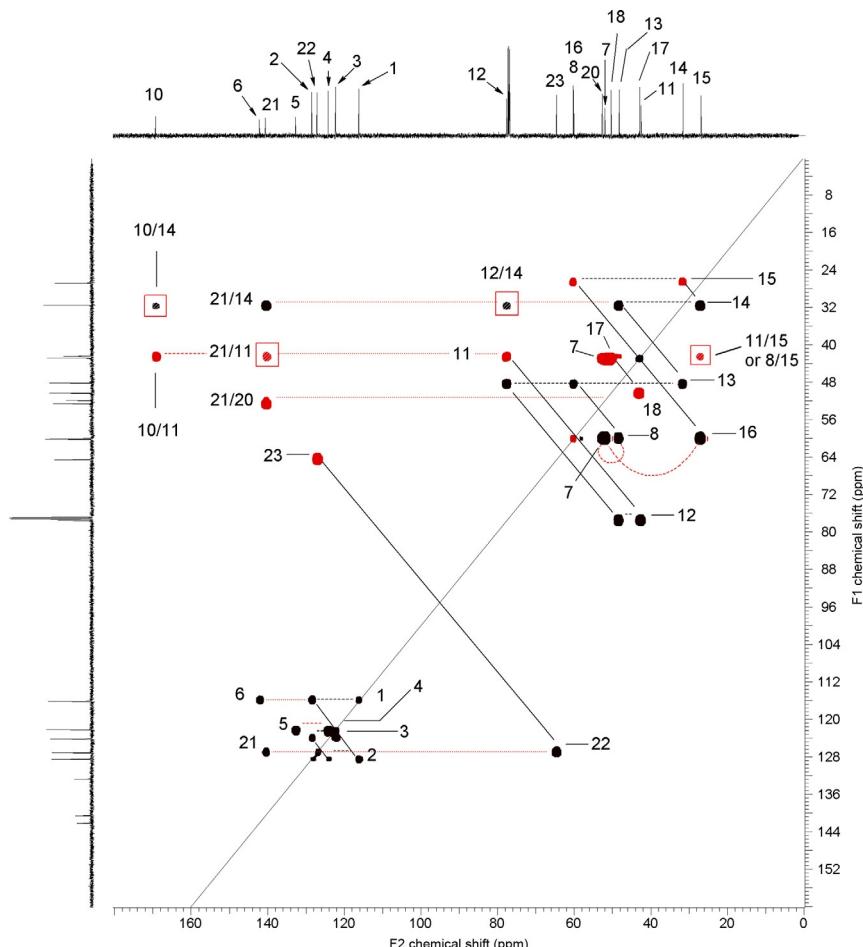
lysine, shikimate and on an extract from the human prostate cancer cell line DU145 [13]. The combination of a regularized H–H 2QF COSY flanked by H–C HSQC to yield a C–C correlation spectrum was pointed out for the ultra high resolution obtained. The spectrum allowed to identify the skeletal structure of individual components from the metabolic mixture or at least fragments thereof. The broad water resonance was successfully removed by moment filtering. The filtering step compared first and second moments of correlation signals. Due to the nature of the parent experiments, quaternary carbon connectivity information was precluded, but it was suggested that inclusion might be possible through replacing one HSQC matrix by a HMBC. The covariance formalism was described as especially suited for the analysis of complex mixtures.

### 5.3. Unsymmetrical and Generalized Indirect Covariance

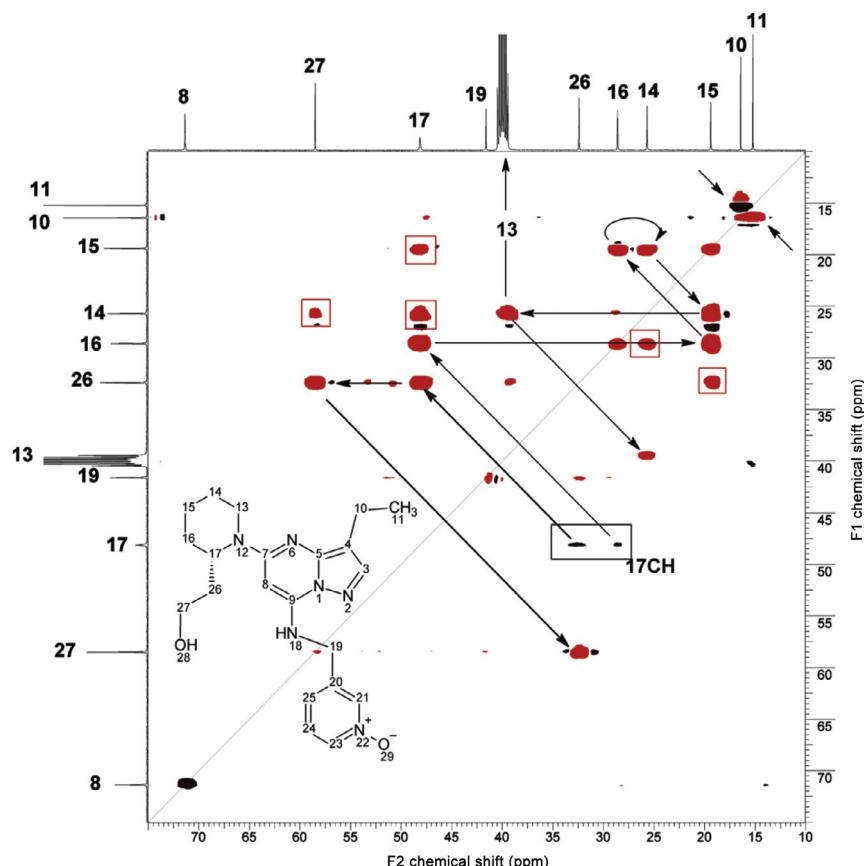
In their recent series of work on structure elucidation and structure verification, Martin *et al.* introduced a concatenated representation of H–C HSQC and 1,1-ADEQUATE spectra based on UIC and GIC [28]. The approach was extended to 1,*n*-ADEQUATE experiments and the replacement of HSQC by HMBC spectra [29,30].

The progress in NMR hardware to support the structure elucidation of limited quantity samples allowed to exploit the ADEQUATE experiment to a greater extent [2,93,94]. The advantage of the ADEQUATE pulse sequence originates from C–C magnetization step transfers that allow to specifically differentiate between  $^2J_{\text{CH}}$  and  $^3J_{\text{CH}}$  bonds which is not possible from HMBC-based experiments. It also includes connectivity information to non-protonated carbon atoms. Thus, multiplicity-edited H–C HSQC and 1,1-ADEQUATE experiments were co-processed to yield a C–C correlation plot [28]. The map was diagonally symmetric in case of adjacent  $J_{\text{CC}}$ -coupled protonated carbons and asymmetric in case of  $J_{\text{CC}}$  between protonated and quaternary carbons. It was emphasized that the latter responses were observed at the  $^{13}\text{C}$  shift of the protonated carbon in the F1 direction and the correlation at the  $^{13}\text{C}$  shift of the quaternary carbon in the F2 dimension, cf. exemplarily for the methylene C11 and the carbonyl C10 in Fig. 5.12. The well-known compound strychnine served as proof-of-principle for the HSQC-1,1-ADEQUATE.

The signal-to-noise ratio of the HSQC-1,1-ADEQUATE was estimated comparable to the parent HSQC. The multiplicity information of the edited HSQC was retained. In a follow-up report, the data requirements



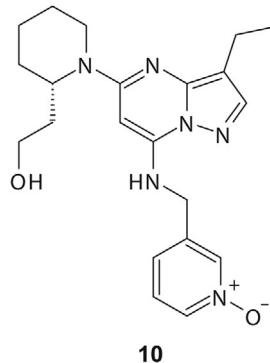
**Figure 5.12** HSQC-1,1-ADEQUATE C–C correlation plot of strychnine **1** from a multiplicity-edited HSQC spectrum and a 60 Hz optimized 1,1-ADEQUATE spectrum by UIC processing. The HSQC-1,1-ADEQUATE correlation plot is diagonally symmetric for signals of protonated carbon correlated via  ${}^1J_{CC}$  with the resonance multiplicity-editing information from the HSQC spectrum retained. Methylene carbon correlations have negative signal intensity (shown in red (gray in print version)) while methine and methyl resonances have positive intensity (shown in black). Carbon–carbon correlations between vicinal protonated carbons are designated by solid diagonal black lines. Correlations between vicinal protonated and non-protonated carbons produce diagonally asymmetric responses at the  ${}^{13}C$  shift of the protonated carbon to which the non-protonated carbon is coupled by  ${}^1J_{CC}$ . Correlations between protonated and non-protonated vicinal carbons are designated by dashed red (gray in print version) lines. Horizontal dashed black lines designate additional correlations from a given carbon. Responses enclosed by red boxes represent ‘spurious’ responses that arise from resonance overlaps. The C10 carbonyl resonance is not shown. Reprinted with permission from Martin et al. [28]. Copyright 2011 John Wiley & Sons, Ltd.



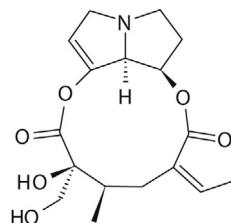
**Figure 5.13** Expansion of the aliphatic region of the 40 Hz HSQC-1,1-ADEQUATE spectrum of Dinaciclib **10**. A convenient starting point is provided by the piperidine methine carbon resonance at  $\sim 47$  ppm enclosed in the black box. Two correlation pathways can be traced beginning from the  $^{17}\text{CH}$  resonance. One coupling pathway defines the pendant  $\beta$ -hydroxy ethyl moiety (C26, C27). The other connectivity pathway sequences and assigns the resonances of the remainder of the piperidine (C16–C13). Responses enclosed in red boxes are artefacts that arise due to resonance overlaps proton spectrum in the multiplicity-edited HSQC and 1,1-ADEQUATE spectra used in the GIC calculation of the HSQC-1,1-ADEQUATE spectrum. Reprinted from Martin et al. [96]. Copyright 2011, with permission from Elsevier.

for minimizing experimental time were investigated [95]. To this purpose, 1,1-ADEQUATE spectra of strychnine were recorded at 500 and 600 MHz using cryoprobes and optimized. Common data array sizes of  $2048 \times 160$  for 1,1-ADEQUATE spectra were acquired. Experimental times ranged from approximately 3 to 19 h. The transformation to HSQC-1,1-ADEQUATE

spectra yielded in all cases a sensitivity enhancement corresponding to the H–C HSQC being the more sensitive experiment. The strength of the method was instructively demonstrated on the cyclin-dependent kinase inhibitor Dinaciclib™ **10** [96]. An illustrative spectrum from this study is presented in Fig. 5.13.

**10**

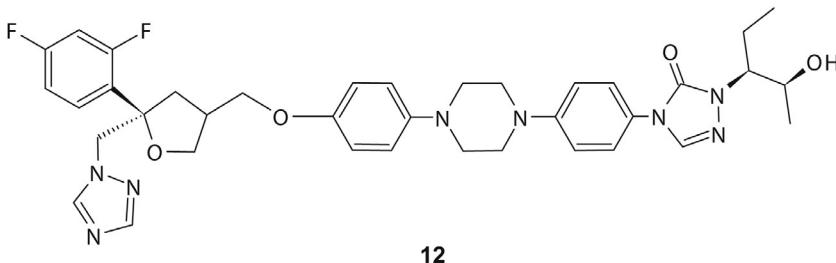
The GIC HSQC-1,1-ADEQUATE was composed from HSQC and 1,1-ADEQUATE, both consisting of  $2048 \times 160$  data points, that were processed to  $1\text{ k} \times 1\text{ k}$  prior to GIC. The result is presented in Fig. 5.13. The complete assignment proceeded smoothly with the help of the C–C correlation scheme and its distinction between  $^2J$  and  $^nJ$  ambiguities. Another example was given for the macrolide retrorsine **11** [33].

**11**

In addition to 1,1-ADEQUATE, 1,*n*-ADEQUATE spectra were incorporated into the HSQC-ADEQUATE correlation map. Although recording times of the latter augmented to 61 h, the access to  $^4J_{\text{CH}}$  correlations, made visible through GIC co-processing with a HSQC, was considered worthwhile. By means of the usual connectivity walk-through known from symmetric COSY or TOCSY spectra, the carbon skeleton could be mapped from the GIC HSQC-1,1-ADEQUATE and HSQC-1,*n*-ADEQUATE. Only adjacent non-protonated carbons and annular heteroatoms interrupted

the connectivity path of the former spectrum. The C–C correlation spectra also led to a reduction of signal overlap because of the higher  $^{13}\text{C}$  dispersion. The sensitivity enhancement of the ADEQUATE information by the GIC transformation was emphasized as a key element for the promotion of the comparatively low sensitivity but high specificity of ADEQUATE experiments. Potential was predicted for manual but also for computer-assisted or automated structure elucidation as a conclusion of the studies summarized here.

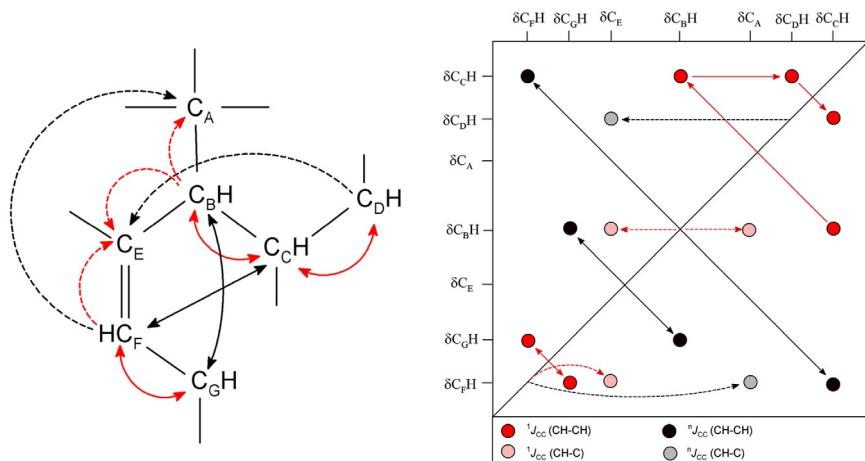
Very recently, the signal assignment of the antifungal agent posaconazole **12** and the structure elucidation of its degradants was reported [97,98]. Both relied heavily on the GIC HSQC-1,1-ADEQUATE derived from multiplicity-edited HSQC and 1,1-ADEQUATE spectra.



The extension of the use of other variants of ADEQUATE combinations through GIC was also tested. In a first study, the 1,1-ADEQUATE was replaced by a 1,*n*-ADEQUATE from a strychnine sample [29]. Experiment times amounted to 48 h on a 600 MHz instrument equipped with a cryo-probe. The information obtained from the HSQC-1,*n*-ADEQUATE spectrum was judged equivalent to that of a  $^4\text{J}_{\text{CH}}$  HMBC. In comparison, a HSQC-1,1-ADEQUATE would be equivalent to a  $^2\text{J}_{\text{CH}}$  HMBC. Although the sensitivity of the 1,*n*-variant is by far inferior to a 1,1-ADEQUATE, the sensitivity increase provided by the HSQC incorporation was sufficient to render the weaker long-range correlations visible. Again the long-range correlations to non-protonated carbons were observed at the F1 chemical shift of the protonated carbon and the F2 chemical shift of the non-protonated carbon in analogy to the HSQC-1,1-ADEQUATE, cf. Fig. 5.13. To access long-range connectivity information, the 1,1-ADEQUATE was combined with a HMBC to yield a HMBC-1,1-ADEUQATE [30]. The result was considered a single-quantum representation of an *n*,1-ADEQUATE, which is the equivalent double-quantum representation. Two disadvantages were found inherent in the latter: its very

low sensitivity testified by only one application report and the tedious interpretation on account of the double-frequency calculations. The proof-of-principle study was conducted on strychnine at 600 MHz using a cryoprobe. As a drawback of the spectrum, it was recognized that the interpretation was not straightforward since magnetization transfer can in principle proceed via many pathways. The  $n,1$ -ADEQUATE starts out with a HMBC step of unknown length  ${}^nJ_{\text{CH}}$ . Yet, the representation in single-quantum frequencies was still seen as an advantage over the corresponding double-quantum spectrum.

Instead of a multiplicity-edited HSQC as in the above experiments, a  ${}^1J_{\text{CC}}$  inverted  $1,n$ -ADEQUATE was covariance transformed with a non-edited HSQC. A  ${}^1J_{\text{CC}}$ -edited HSQC- $1,n$ -ADEQUATE resulted from the GIC formalism [31]. For the inversion of the  ${}^1J_{\text{CC}}$  correlations, 19 h of experiment time divided on 160 increments in the  $t_1$  domain were spent on a 600 MHz NMR instrument equipped with a 1.7 mm cryoprobe. The sample tube contained 500  $\mu\text{g}$  of strychnine. A schematic representation of the  ${}^1J_{\text{CC}}$ -edited HSQC- $1,1$ -ADEQUATE is given in Fig. 5.14.



**Figure 5.14** Schematic representation (right) of the  ${}^1J_{\text{CC}}$ -edited HSQC- $1,n$ -ADEQUATE spectrum that can be calculated from an inverted  ${}^1J_{\text{CC}}$   $1,n$ -ADEQUATE and unedited HSQC spectrum via covariance processing. The four types of correlations possible are shown in the legend beneath the schematic of the contour plot. Parenthetical information indicates correlations between pairs of protonated carbons (CH-CH) or between a protonated and non-protonated carbon (CH-C) via either  ${}^1J_{\text{CC}}$  or  ${}^2J_{\text{CC}}$  coupling pathways designated of the structure on the left. Reprinted with permission from Martin et al. [31]. Copyright 2012 John Wiley & Sons, Ltd.

The spectral interpretation with reference to Fig. 5.14 was described in the following way.

1. The correlations between adjacent protonated carbons are indicated as double-headed solid red (gray in print version) arrows. Diagonally symmetric signals with negative, or inverted, phases occur in the spectrum.
2. Correlations between adjacent protonated and non-protonated carbons are connected by single-headed dashed red (gray in print version) arrows. The signals are diagonally asymmetric.
3. Correlations between long-range coupled protonated carbons are denoted by double-headed solid black arrows. Diagonally symmetric signals with positive phase are observed.
4. Correlations between long-range coupled protonated and non-protonated carbons are denoted by dashed, single-headed black arrows. Their signals are diagonally asymmetric.

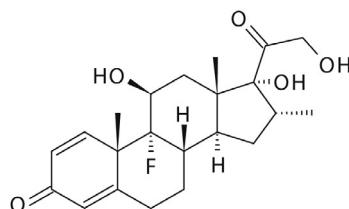
The corresponding signal types were found in the spectrum of strychnine. With the exception of correlations between quaternary carbons the complete carbon skeleton of a molecule could be traced out by means of a GIC  $^1J_{\text{CC}}$ -edited HSQC- $1,n$ -ADEQUATE spectrum [31].

As could be recognized from the review of the more recent applications, the use of UIC has been abandoned in favour of GIC. Since GIC treatment would lead to qualitatively and semi-quantitatively similar results, the forthcoming section summarizes the original achievements based on UIC with reference to future utilization of the correlation maps or spectra, which will then be computed according to GIC.

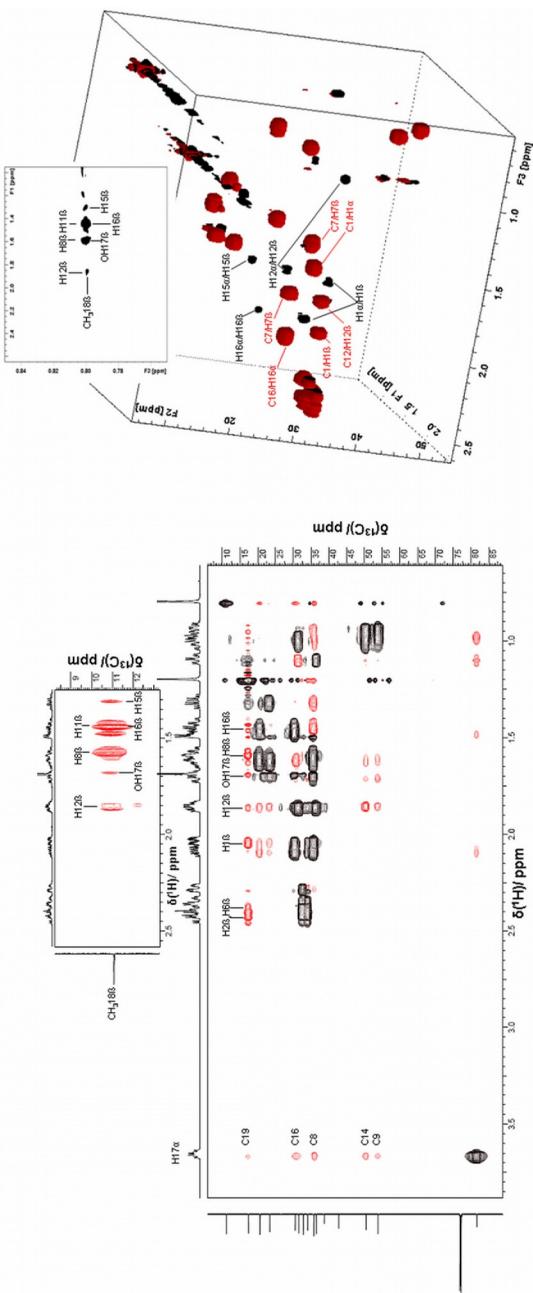
To illustrate the first application of GIC, Snyder *et al.* transformed H–H TOCSY and H–C HMBC spectra to GIC HMBC–TOCSY correlation maps according to the formalism in Eq. (5.27) to Eq. (5.31) [14]. Seven metabolites were selected, D-carnitine, D-glucose, L-glutamine, L-histidine, L-lysine, myo-inositol, shikimic acid and the MDM2-binding p53 peptide construct with the sequence ETFSDLWKLLPEN. The TOCSY spectra were acquired with  $2048 \times 1024$  or  $2048 \times 256$  data points, the H–C HMBC with  $2048 \times 1024$  points at 800 MHz using cryoprobe technology. In addition to the generalization of the indirect covariance formalism, a method for artefact recognition was derived from the mathematical description. While the square root operation on the covariance map was already found to lower the number of artefacts by a factor of three, it was demonstrated that artefacts could be recognized if the peak intensity of a tentatively false correlation was plotted as a function of  $\lambda$  in comparison with the trend of true signals. Furthermore, the GIC HMBC–TOCSY accounted for

the visualization of the connectivity information of quaternary carbons, for example, carbonyl and carboxyls, which proved very suitable for peptides. As indicated above, the advent of GIC and its advantageous effect on artefact suppression of UIC subspectra slowly replaced the sole use of UIC. Probably, the last broad study relying exclusively on UIC, albeit in combination with indirect covariance processing, was performed by Aspers *et al.* [24].

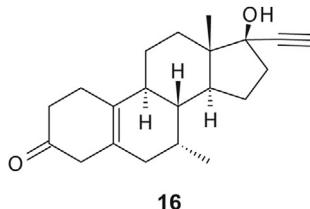
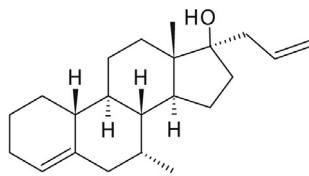
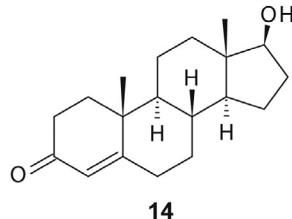
Four steroidal compounds, dexamethasone **13**, testosterone **14**, allylestrenol **15**, tibolone **16**, exhibiting different levels of spectral crowding in the proton as well as in the carbon domain were analysed by means of H–H TOCSY with different mixing times, H–H NOESY, H–C HSQC, 2D H–C HSQC–TOCSY and 3D H–C NOESY–HSQC. Besides C–C TOCSY correlation maps, C–C NOESY maps were generated from UIC HSQC–NOESY spectra. The spectra generated by covariance transformations were compared to their experimental counterparts, except for the C–C correlations, with respect to signal-to-noise ratios, experiment time, number of true and false correlations. In particular, the information gain by multiplicity-edited or phase-sensitive experiments was discussed, since signal phases are conserved through UIC covariance processing. All results were considered in relation to the spectral congestion of the compounds studied. For application in routine laboratories, it was suggested to pay attention to experimental parameter optimization, for example, to minimize  $t_1$ -noise, to phase-sensitivity and to the signal density, for example, caused by longer TOCSY mixing times, before potential covariance treatment. The authors found that with increasing spectral congestion the level of confidence for structure elucidation would be lowered. As an example for a significant reduction in experiment time, a 2D H–C HSQC–NOESY obtained by transforming two rather high-sensitivity experiments, that is, H–H NOESY and H–C HSQC, is illustrated in Fig. 5.15 together with its experimental equivalent in form of a 3D HSQC–NOESY. The insets in the spectra display the slices necessary to solve a specific assignment problem of testosterone **14**.



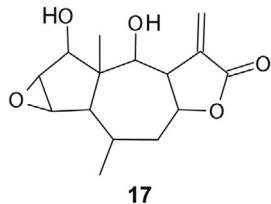
**13**



**Figure 5.15** Comparison of H–C HSQC–NOESY spectra of testosterone **14**. UIC 2D HSQC–NOESY (left), 3D HSQC–NOESY (right). Correlations used for the assignment of H16 $\alpha$  and H16 $\beta$  resonances are indicated in the zoom regions. The F1–F3 plane (H–H plane) of the 3D cube is shown at carbon chemical shift of the 18-methyl resonance. *Reprinted from Aspers et al. [24]. Copyright 2011, with permission from Elsevier.*

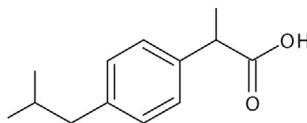


The earlier pioneering and explorative studies dealt individually with the aspects of the generation of hyphenated correlations spectra, the recognition or elimination of artefacts and the significant increase in sensitivity or a reduction of experimental time as compared to the full acquisition of the hyphenated experiments. Blinov *et al.* computed H–C HSQC–TOCSY and HSQC–COSY spectra from the corresponding component spectra of the sesquiterpene lactone autumnolide **17** [19].



While the hyphenated experiments were recorded in 16 h, the component spectra were acquired between 10 min and 1 h each, covariance

processing proceeded within seconds on a then standard personal computer. From the UIC-hyphenated spectra, C–C correlations were calculated as well. Since their equivalent lies in a C–C INADEQUATE or H–C ADEQUATE experiments, no comparison with respect to experiment time was given. However, a 10-fold signal-to-noise enhancement was reported for the UIC-hyphenated spectra, although the calculations were performed without taking non-linear noise levels into account [36]. A UIC HSQC–COSY was also applied to help disentangle potentially overlapped connectivity information of strychnine **1** [20]. The individual parent spectra consisted of  $1024 \times 256$  data points, COSY, and  $1024 \times 96$  data points, HSQC, but were zero filled to  $2048 \times 512$ . Through UIC processing, a HSQC–COSY could be obtained that was found to provide a 22-fold improvement with respect to sensitivity. The UIC H–C HSQC–NOESY was subsequently introduced, investigating a sample of ibuprofen **18** [23].



**18**

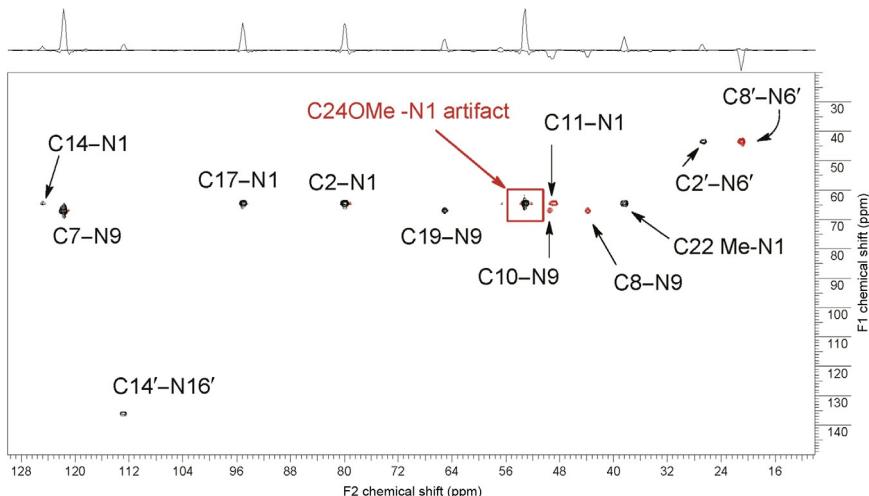
Whereas an equivalent experiment was performed within roughly 45 h, HSQC and NOESY were acquired within 30 min and 3–7 h resulting in UIC-combined HSQC–NOESY spectra presenting the same information content. As a common problem in covariance processed spectra, the appearance of false signals originating from signal overlap in the component or parent spectra was evenly observed in hyphenated spectral representations [43]. Therefore, a study on the polycyclic aromatic compound **4** was devoted to find a general means for the identification of artefacts. Type I and II artefacts were thus identified in hyphenated UIC spectra. As described above, the first type referred to pairs of direct or relayed correlations with a negative phase, if true correlations had opposite phase. The latter arose between direct and relayed responses of different spin systems and had positive phase, thus making them identifiable only through analysis. It was proposed to subject the HSQC spectrum being usually one of the parent spectra to indirect covariance transformation. Off-diagonal elements could only stem from proton overlap in this C–C correlation map, therefore rendering the recognition of possible artefacts in the hyphenated spectrum possible [43].

The last reports discussed the concatenation of homonuclear with heteronuclear correlation spectra. Yet, the combination of two similar

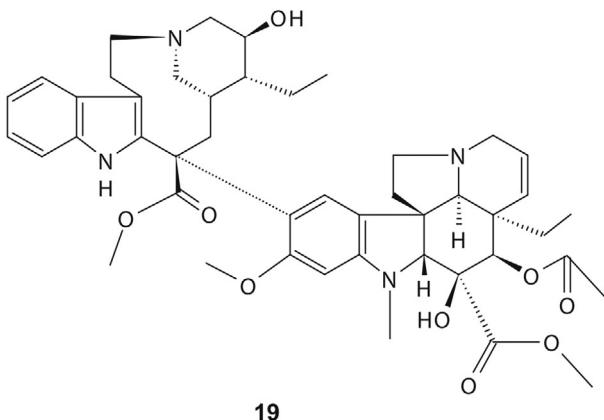
heteronuclear correlations was carried out for H–C HSQC and H–C HMBC spectra of strychnine [22]. From the UIC processing, a long-range C–C correlation map resulted, which can be regarded analogous to *n*,<sup>1</sup>-, 1,*n*-, and *m,n*-ADEQUATE spectra. Since both component spectra provide a much higher sensitivity than a comparable ADEQUATE experiment, the long-range C–C correlation was suggested as a promising alternative. Another variant of a long-range C–C correlation was calculated from a H–C IDR HSQC–TOCSY and a H–C HMBC of strychnine to yield a HSQC–TOCSY–HMBC [25]. The spectrum revealed two- and four-bond correlations that appeared much weaker in the HMBC. While in terms of signal-to-noise the UIC spectrum seemed superior, the HMBC was said advantageous where carbon signals were not resolved but the corresponding proton signals were as those correlations did not occur in the HSQC–TOCSY–HMBC. Nevertheless, the spectrum displayed very many correlations that presented limitations due to <sup>13</sup>C spectral congestion and overlap.

In the same direction of research, UIC was used to create representations of correlation between nuclei that have a low gyromagnetic ratio and limit the access to corresponding spectra if no stable-isotope labelling is applied to the compounds under investigation. In the domain of solution-state NMR, these correlations might be referred to as exotic. As small molecules are often constituted of protons, carbon and nitrogen atoms—oxygen shall not be considered at all due to its mixture of very unfavourable NMR properties of <sup>17</sup>O—the correlation between <sup>13</sup>C and <sup>15</sup>N was among the target spectra. Hence, a multiplicity-edited H–C HSQC and a H–N HMBC were UIC transformed into a C–N correlation, referred to as C–N HSQC–HMBC [42]. This otherwise inaccessible spectrum of a few milligrammes of strychnine was constructed after a total recording time of less than 4 h for both component spectra. The spectral representation was predicted to be of interest for the structure elucidation of pharmaceutical actives, agrochemicals and alkaloids due to their elevated nitrogen content.

After the feasibility study on strychnine, the bis-alkaloid vinblastine **19** was investigated by C–N HSQC–HMBC obtained again from H–C HSQC and H–N HMBC [27]. For the prediction and recognition of artefacts from spectra overlap, Martin *et al.* applied the approach based on indirect covariance of the parent HSQC, and aromatic solvent induced shift, which proved successful. The authors stated that no new information was created by UIC but a favourable visualization scheme was used. Figure 5.16 presents an example of a C–N correlation.

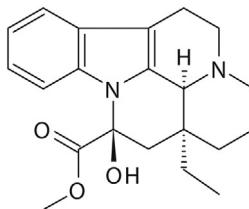
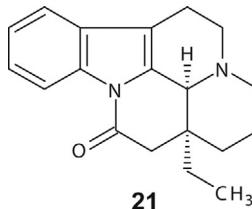


**Figure 5.16** C–N HSQC–HMBC spectrum showing response assignments of **19**. The correlation response from the C24 O-methyl group to N-1 is an artefact. The overlap of the H-2 methine and the C24 O-methyl resonances could be recognized in the parent multiplicity-edited H–C HSQC. This type of overlap can give rise to artefact responses in UIC processed heteronuclear correlation data matrices. *Reprinted with permission from Martin et al. [27]. Copyright 2007 American Chemical Society.*

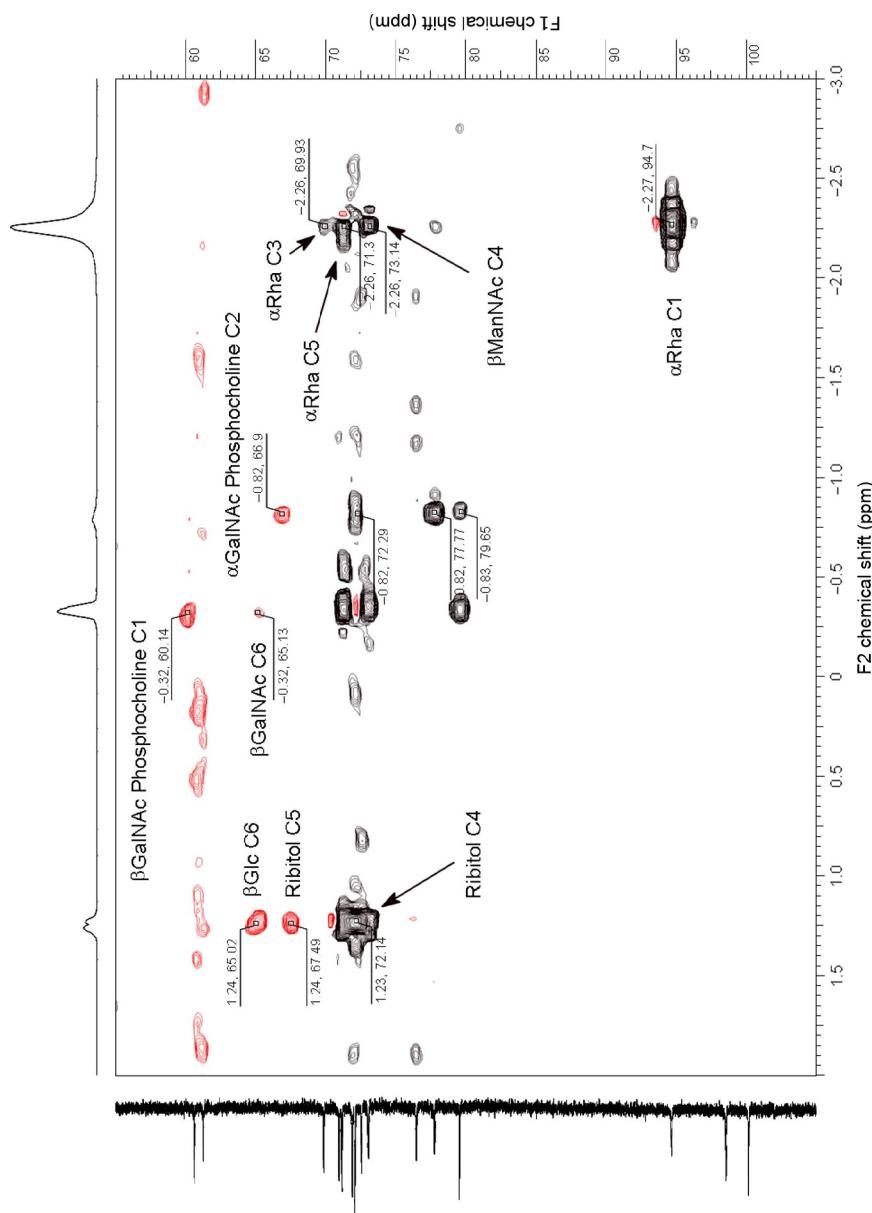


Another alkaloid example was studied by UIC C–N correlation spectroscopy [41]. Spectra of the indole alkaloid vincamine **20** were acquired as H–C HMBC with  $2048 \times 160$  data points and H–N IMPEACH–HMBC with  $1024 \times 96$  data points. The calculated correlation map was called C–N HMBC–IMPEACH. It provided a high number of correlations, which were considered to hamper structure elucidation more than to enhance.

Applicability was nevertheless suggested for structure confirmation. The spectrum was also compared to a C–N HSQC–IMPEACH correlation from H–C HSQC H–N IMPEACH–HMBC spectra. The plot presented  $^2J_{\text{CN}}$  and  $^3J_{\text{CN}}$  correlations from  $^3J_{\text{NH}}$  and  $^4J_{\text{NH}}$  pathways. Due to its reduced number of traceable signals, it was judged of broader usability. The C–N HSQC–IMPEACH was successfully employed in the signal assignment of (–)-eburnamonine **21** [40].

**20**

Recently, a C–P correlation spectrum was introduced [26]. The spectrum was computed from H–P HMBC and H–C HSQC spectra recorded at 700 MHz using cryoprobe technology on a sample of the polysaccharide vaccine STF19 from *Streptococcus pneumoniae* serotype F19 (**PS**) containing the contaminant teichoric acid (**c-PS**). In polysaccharide analysis, a challenging task consists of the site specific identification of phosphoesters in mixtures of saccharides. In proton-detected experiments, strong resonance overlap poses often problems. Therefore, C–P correlations were used to circumvent the problem. With the help of GIC, the sensitivity enhancement accounted for the detection of  $^{3-5}J_{\text{HP}}$  couplings. As the result, all phosphoester moieties could be unambiguously assigned. A representative spectrum C–P spectrum of the mixture is given in Fig. 5.17.



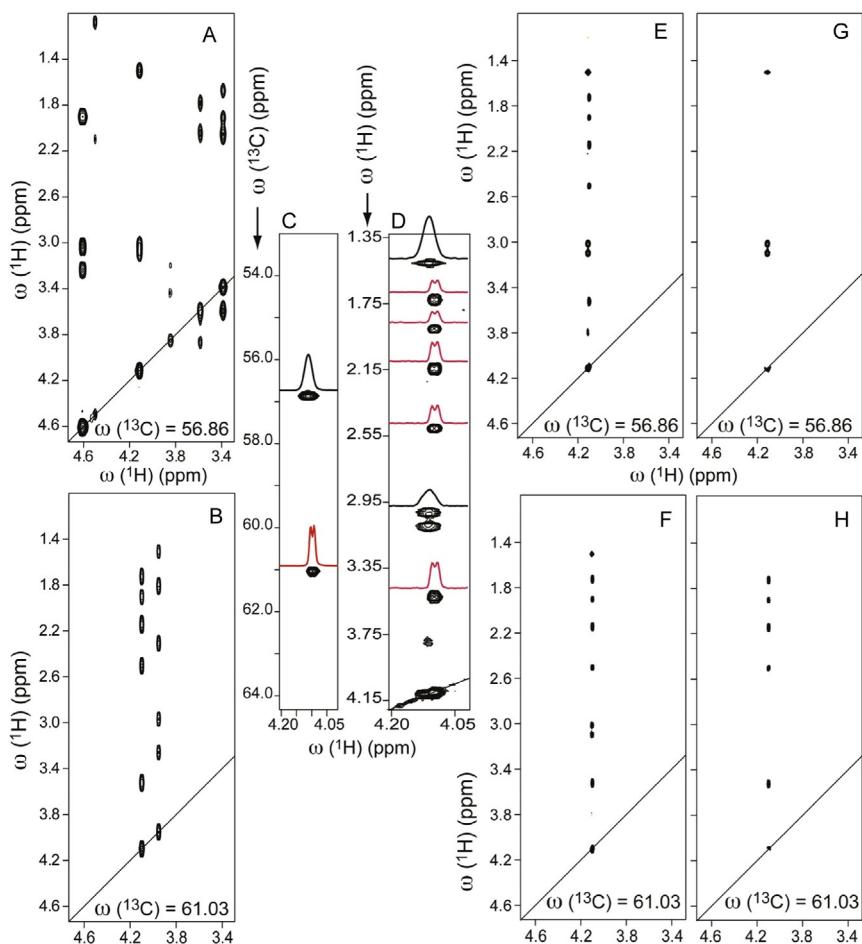
**Figure 5.17** Expansion of the P-C covariance spectrum. Assignments of phosphodiester from **PS** ( $\delta\text{P} = -2.27$  ppm) and the three phosphodiesters from **c-PS** ( $\delta\text{P} = 1.26$ ,  $-0.34$ , and  $-0.81$  ppm). Unambiguous assignments from this calculated spectrum are labelled. With kind permission from Springer Science + Business Media: Zartler and Martin [26].

## 5.4. Multidimensional Covariance

### 5.4.1 Triple Rank

3R covariance spectroscopy as a representative of multidimensional covariance NMR was applied to the cyclic decapeptide antamanide [15]. Two 3R HSQC–TOCSY spectra were computed from a H–C HSQC and H–H TOCSY spectra with 25 and 90 ms mixing times, each consisting of  $1536 \times 900$  data points. The spectra were zero filled to  $2\text{ k} \times 1\text{ k}$  points prior to covariance processing. The 3R spectrum thus generated was composed of 1024  $^{13}\text{C}$ -edited TOCSY planes which means ultra high resolution as individual proton spin systems were sorted into different planes. The spectra obtained were compared to 3D HSQC–TOCSY obtained from acquiring  $2048\text{ (}^1\text{H direct)} \times 128\text{ (}^{13}\text{C)} \times 256\text{ (}^1\text{H)}$  points and zero filling to  $2\text{ k} \times 1\text{ k} \times 1\text{ k}$  and Fourier transformation. The 3R spectrum provided much better resolution in the carbon-dimension. Comparative spectra from the study are presented in Fig. 5.18. The size of the spectrum was reported as 8.1 GB, but did not cause any problem to modern computer equipment. In combination with covariance transformation, spectral filtering based on signal mean positions and their line widths was performed. As a result, 22 of the 25 overlapping proton resonances of antamanide could be successfully disambiguated.

Another 3R covariance study was dedicated to the analysis of complex mixtures [99]. A mixture of carnitine, alanine, isoleucine, ornithine, arginine, lysine, shikimate and glutamate were analysed through 3R HSQC–TOCSY as well as an extract from the *E. coli* BL21 (DE3) strain. The covariance spectra were deconvolved for mixture component identification through a consensus plane extraction and clustering (nD DeCoDeC) approach. Provided that each component of the mixture had one resolved TOCSY cross-peak, the consensus approach was used to identify common peaks from cross-sections along both dimensions of each covariance TOCSY cross-peak. The corresponding 1D TOCSY traces were hence extracted and clustered. This approach was adopted to H–C HSQC planes. A third strategy extracted pure 2D H–C HSQC spectra of the individual mixture components from the 3R HSQC–TOCSY spectrum by means of a 2D version of the consensus algorithm. A similarity measure between pairs of traces was defined, which permitted clustering. The clustering results were displayed as dendograms. While the 2D DeCoDeC allowed the determination of 1D  $^1\text{H}$  spectra, the 3D DeCoDeC version permitted to extract H–C HSQC spectra of individual mixture constituents.



**Figure 5.18** Performance of triple-rank correlation method with moment filtering for antamanide. For the two overlapping HR resonances of Pro3 (red (gray in print version)) and Phe5 (black), (C and D) show 2D HSQC and TOCSY ( $\tau_m = 90$  ms) strip plots; (A and B) show 3D FT strip plots at constant  $^{13}\text{C}$  frequency; (E and F) show 3R planes before filtering; and (G and H) show those after filtering. Peak cross-sections in (C) and (D) demonstrate differential  $^1\text{H}$  line positions and line widths that permit identification of false peaks in (E) and (F) and their successful suppression in (G) and (H) based on moment filtering. Reprinted with Permission from Bingol et al. [15]. Copyright 2010 American Chemical Society.

#### 5.4.2 4D NMR

The use of multidimensional NMR spectroscopy is more common to protein NMR than to small molecule NMR. As our focus lays on small-molecule applications, we will not extend our review into the field of

protein NMR. Since 4D covariance analysis in the form discussed above was described by Snyder *et al.*, two of their applications on protein samples will be summarized here [16,100]. The compound under investigation was  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled ubiquitin. An 800 MHz spectrometer equipped with a cryoprobe was used. A  $^{13}\text{C}$ -edited 4D NOESY of dimensionality  $955 \times 8 \times 32 \times 96$  points,  $^1\text{H}$  acquisition,  $^{13}\text{C}$ ,  $^1\text{H}$ ,  $^{13}\text{C}$ , respectively, was recorded within 130 h. A shared evolution N/C-edited NOESY was acquired in 96 h. The former 4D NOESY experiment was symmetric and could therefore be subjected to covariance transformation following the plane-of-planes strategy, cf. above. The  $^{13}\text{C}$ -HSQC–NOESY– $^{15}\text{N}$ -HSQC was not symmetric, but the 4D covariance scheme could be generalized such that NOE distance information between H–N and H–C groups could be extracted from the N/C-edited NOESY. The transformed representations also allowed the observation of additional peaks with long-range structural information. The 4D covariance was applied after Fourier transformation of both the direct proton dimension and the dimension of the heteroatoms attached to the protons of the direct dimension. The concept of donor/acceptor planes and pairs was applied. The covariance processing led as usual to a substantial resolution enhancement.

We consider covariance NMR of small molecules as the focal point of this chapter. The numerous accounts on protein studies, where full covariance processing, partial covariance processing or covariance analyses of constraints or protein dynamics is used, are beyond the scope of this overview. Just for illustration purposes of the different aspects that can be addressed with covariance concepts, seven reports from the past decade are arbitrarily selected [101–107].

## 5.5. Synchronous and Asynchronous Spectra

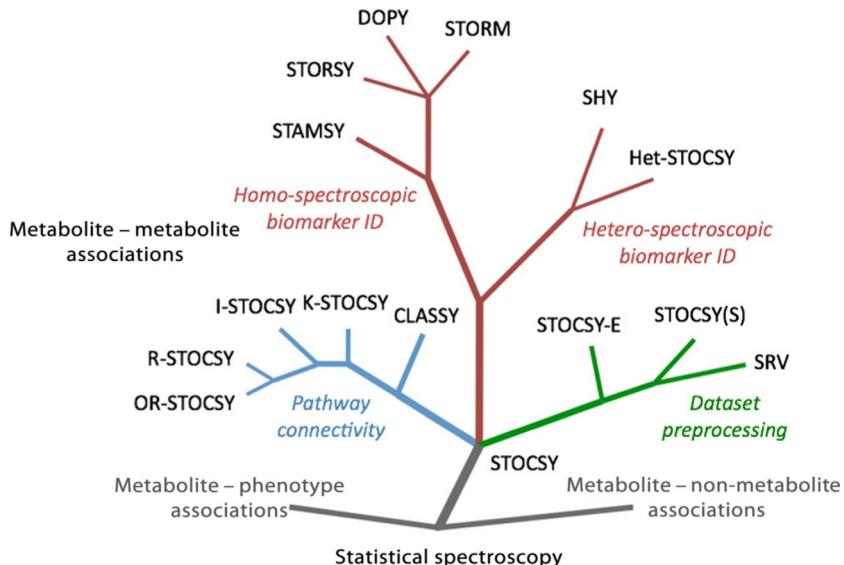
The applications of covariance processing in the general field of spectroscopy are regularly reviewed [3,4,48,108,109]. Trends of developments in NMR among them covariance NMR are also published periodically [110–115]. In the context of this review, three reports were selected within this section. Kirwan *et al.* applied generalized 2D covariance NMR to analyse the fermentation process of wine [44,116]. They selected Shiraz wine from the McLaren Vale in Australia. They recorded 1D  $^1\text{H}$  NMR spectra of 26 samples taken at successive days during the fermentation of wine. The indirect dimension was thus represented by the series of samples. In analogy to metabolomics studies, spectral alignment, pre-processing to a uniform

line width for all samples, normalization to a standard and the binning of spectral regions were found prerequisites for successful generation of synchronous and asynchronous correlation maps. While the synchronous maps were less prone to small chemical shift and line-width variations, the asynchronous matrices were found very sensitive. The temporal relationship, that is, the co-variation during fermentation, could be interpreted from the covariance maps. The correlation between species, which in other words is the conversion process, could be recognized from the off-diagonal signals in the synchronous matrix, which corresponds to the direct covariance map. The asynchronous map, which would be equivalent to an indirect covariance transformation albeit on the sample dimension, carried the information on the relative rate of change of the correlated species in agreement with Noda's rules. The results derived on the fermentation process were found congruent with previous knowledge, cf. references cited in the described study [116]. A detailed description of the effects of line width and chemical shift variables by Sasic preceded the reports by Kirwan *et al.* [46]. In his metabonomics study of rat urine samples as a test for vasculitis, butterfly-like signal shapes were found to be a result of wandering peak positions. The lack of uniform pre-processing led to numerous artefacts and problems that severely hampered spectral interpretation in contrast to the wine study.

## 5.6. Statistical Analysis of NMR Spectra in the Field of Metabolomics

The use of statistical, mathematical tools for the transformation and analysis of NMR spectra with applications to metabolomics seemed having taken a path independent of the covariance processing of spectra used in the structure elucidation of isolated small compounds. The NMR methodology to address the issue of unraveling complex mixtures in biological fluids, such as urine or blood, started out as STOCSY. The STOCSY-type experiments were employed to assign chemical structures, prepare spectroscopic data sets for pattern recognition analysis, and identify co-regulation between metabolites. From the roots of this strategy over the branches to its full exploitation in the field of medicine and systems biology, the developments are intimately related to Nicholson and coworkers [117]. The extent of methods available today can be recognized from Fig. 5.19.

With respect to covariance NMR, STOCSY themes can be interpreted as a form of direct covariance, where the direct dimension consists of the proton chemical shift and the indirect dimension is represented by sample variations.



**Figure 5.19** Statistical correlation spectroscopy family tree. Branches in the tree show different statistical spectroscopic tools and indicate the major applications of these tools in the field of metabolomics. Reprinted with permission from Robinette et al. [118]. Copyright 2013 American Chemical Society.

From the perspective of generalized covariance spectroscopy, it may be regarded as the synchronous map that correlates variations in the sample domain. Since the variations are not caused by magnetization transfer, correlations do occur between multiple spin systems thus connecting species, for example, metabolites that may be chemically or biologically related. These relations may lead to the identification of metabolite–metabolite associations, metabolite–non-metabolite associations, and metabolite–phenotype associations of metabolic and ultimately of regulation pathways. When different types of spectroscopy or spectrometry are applied and correlated through the sample dimension, the method was termed SHY and was reviewed in [Section 2.7](#). From another point of view, STOCSY approaches could be subsummarized under the topic mixture deconvolution, since the elucidation of complex mixtures containing biogenic small molecules is one of its targets. A forward-looking conspectus pursuing the comparison of covariance for deconvolution and STOCSY techniques was recently published [\[119\]](#).

The application of STOCSY-derived methods to metabolomics research was reported in numerous studies. Accompanied by the medicinal relevance, an intensively studied field emerged. From the original

STOCSY approach, a range of variants was developed, as already indicated in Fig. 5.19. Some of the variants rely on different experiments or even other spectroscopic methods, while some variants comprise an alternative analysis method. A brief overview of the key aspects of the methods, their application and information, that can be obtained from the individual methods, is compiled in Table 5.3 [118]. We consider the coverage of the STOCSY literature beyond the scope of this chapter. Instead, we suggest the following works for introductory reading on the topic [117–119,135,136].

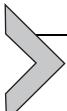
## 5.7. Other Covariance Applications in NMR Spectroscopy

Covariance matrix calculation and analysis were also described for completely other purposes than spectral processing. Covariance matrix residuals were used as a tool to diagnose experimental artefacts in DOSY NMR data [137] or to improve the deconvolution performance in cases of signal congestion [138]. In a one-step many parameter optimization scheme, covariance matrices were contained in an algorithm for the optimization of phase-modulated homonuclear decoupling pulse shapes in solid-state NMR [139]. Changing the field of application again, covariance calculations were implemented into evolutionary strategies to determine dipolar couplings observed in  $^1\text{H}$  NMR spectra of *n*-butane dissolved in a liquid crystal [140]. The same group reported incorporated the covariance matrix adaption evolutionary strategy into a molecular dynamics prediction of the NMR spectra of *n*-pentane as solute in the nematic phase of a liquid crystal [141]. To achieve an enhancement of the signal-to-noise ratio, the phase correlation between an NMR signal and the excitation pulse was analysed by the covariance formalism [142]. The key idea was to obtain the phase of the NMR spectrum at each frequency point and to analyse the phase correlation with the radio frequency phase. This approach was referred to as phase-covariance analysis.

It should be understood that the covariance matrix is first of all per definition a mathematical or statistical representation of variations. The physical interpretation of its values is derived from the data or values that are analysed. The application of the mathematical procedure is thus only limited by the imagination of the researcher. It is hence not unlikely that some beautiful utilizations of covariance analysis, especially if not explicitly emphasized for searches, are not mentioned at this point. We offer our apologies to those authors.

**Table 5.3** Statistical correlation spectroscopy tools and their applications in the field of metabolomics

Statistical correlation spectroscopic method	Purpose/field of application
Statistical total correlation spectroscopy (STOCSY) $^1\text{H}/^1\text{H}$ NMR	Biomarker identification through correlations in 1D NMR [120,121]
Subset optimization by reference matching (STORM) $^1\text{H}/^1\text{H}$ NMR	Identification of structural correlations from minor signals in large NMR data sets [122]
Diffusion ordered projection spectroscopy (DOPY) $^1\text{H}/^1\text{H}$ DOSY NMR	Resolution improvement of correlated signals [123]
Statistical total regression spectroscopy (STORSY) $^1\text{H}/^1\text{H}$ J-res NMR	Identification of stoichiometric relationships between NMR signals [124]
Statistical heterospectroscopy (SHY) $^1\text{H}$ NMR/LC-MS	Correlation between NMR and MS for identification [52]
Heteronuclear STOCSY (Het-STOCSY) $^1\text{H}/^{19}\text{F}$ , $^1\text{H}/^{31}\text{P}$ NMR	Correlation of heteronuclear NMR signals [125,126]
Cluster analysis statistical spectroscopy (CLASSY) $^1\text{H}/^1\text{H}$ NMR	Distinction between structural and pathway correlations and identification of metabolic perturbations [127]
Kinetic STOCSY (K-STOCSY) $^1\text{H}/^1\text{H}$ NMR	Sequence monitoring of biotransformations on short time scales [128]
Iterative STOCSY (I-STOCSY) $^1\text{H}/^1\text{H}$ NMR	Visualization of network of pathways [129]
Recoupled STOCSY (R-STOCSY) $^1\text{H}/^1\text{H}$ NMR	Identification of sets of pathway relationships in perturbed metabolic pathways [130]
Orthogonal filtered recoupled STOCSY (OR-STOCSY) $^1\text{H}/^1\text{H}$ NMR	Identification of metabolites in weakly perturbed networks [131]
STOCSY-editing (STOCSY-E) $^1\text{H}/^1\text{H}$ NMR	Removal of exogenous compound signals [132]
STOCSY-scaling (STOCSY(S)) $^1\text{H}/^1\text{H}$ NMR	Scaling of statistical variance of compound signals [133]
Statistical recoupling of variables (SRV) $^1\text{H}/^1\text{H}$ NMR	Adaptive binning of signals to reduce dimensionality [134]



## 6. CONCLUSION

Covariance NMR has established itself as a valuable tool in the ranks of NMR methodologies. As direct and indirect covariance of 2-dimensional NMR data sets, it can replace the second Fourier transformation, whereas as unsymmetrical and GIC it follows Fourier transformation. Based on matrix algebra and statistical mathematics, covariance transformations were extended to doubly indirect and 3- and 4-dimensional covariance NMR. Since the parent or component data arrays can originate from any type of 2D NMR experiment, 2D NMR spectra were co-processed to yield C–N and C–P correlations of non-isotopically labelled small molecules. Further heterospectroscopic covariance was used to concatenate NMR and MS data allowing to allocate the information of both to a compound.

The combination of NUS schemes appeared very promising, as it accounted for a significant gain in experimental time. Covariance processing proved robust for use with limited, sparsely sampled data sets in contrast to Fourier transformation. Nevertheless, care must be taken with respect to sampling in the range of the Nyquist frequency. No loss in resolution was observed, either, due to the projection or mapping of the acquisition dimension onto the sparsely sampled dimension. Application examples were reported, in particular, for solid-state NMR.

Next to exotic correlation spectra, such as C–N or C–P, covariance NMR was found especially suitable for the generation of hyphenated experiments such as HSQC–TOCSY and HSQC–NOESY. Recent studies heavily relied on HSQC–ADEQUATE correlations for structure elucidation and signal assignment, obtained through GIC processing. Those investigations were motivated by the selectivity of the ADEQUATE experiment to distinguish between  $^1J_{CC}$  and  $^2J_{CC}$  couplings. The use of the 1,1-ADEQUATE was extended to 1, $n$ - and  $m,n$ -ADEQUATE experiments also in combination with  $^1J_{CC}$  editing and a HMBC co-processing. The resulting spectrum in all cases is a C–C correlation representation that allows to map C–C connectivity in the same way as a COSY or TOCSY is interpreted.

Common to the covariance transformations, sensitivity enhancement or experiment time reduction were observed, since in the case of combination spectra the sensitivity of the resulting correlation map would approach the sensitivity of the component spectra. For the HSQC-1,1-ADEQUATE, the signal-to-noise ratio was reported to be nearly as high as for the HSQC itself. The HSQC might hence be used to enhance the signal-to-noise ratio of the

ADEQUATE experiment. Furthermore, resolution improvement was also stated for direct covariance processing, for transposing correlations into a larger heteronuclear spectral dispersion or for creating higher rank spectra. Covariance processing methods were described, that allowed the removal of residual water resonances or to apply individual filtering or thresholding for different spectral regions.

In comparison to other spectroscopic techniques, where generalized covariance was performed to obtain synchronous and asynchronous correlation maps and where both correlation maps were interpreted, the asynchronous map was rarely exploited if considered at all for NMR purposes. Yet, the synchronous map as an equivalent to the direct covariance spectrum served to correlate species in different samples for a few studies. In contrast, analysis of the sample variation by statistical total correlation NMR, STOCSY, has become a corner stone of metabolomics investigations; this field was considered beyond the scope of this chapter, hence only the current variants of STOCSY and their purposes were briefly presented.

With respect to mixture analysis, for which STOCSY was used as well, successful approaches like indirect covariance TOCSY and pure-shift Zanger–Sterk or constant-time indirect covariance TOCSY were collected.

Last but not least, software solutions, that allow covariance NMR processing at various levels of complexity, were compiled together with their commercial or not-for-profit providers.

## ACKNOWLEDGEMENTS

The authors thank M. Voigt, M.Sc., and N. Nöthling, B.Sc., for their help in preparing the chapter. Prof. Dr. K. Graf is thankfully acknowledged for fruitful discussions on matrix algebra. The authors are grateful to Prof. Dr. I. Noda for providing two of his recent manuscripts prior to publication. They are also indebted to Dr. Gary E. Martin for his insightful comments.

## REFERENCES

- [1] D.A. Snyder, R. Bruschweiler, Multidimensional correlation spectroscopy by covariance NMR, in: G.A. Morris, J.W. Emsley (Eds.), *Multidimens. NMR Methods in the Solution State*, John Wiley & Sons Ltd., Chichester, UK, 2010, pp. 97–105
- [2] G.E. Martin, Using 1,1- and 1, n-ADEQUATE 2D NMR data in structure elucidation protocols, *Annu. Rep. NMR Spectrosc.* 74 (2011) 215.
- [3] I. Noda, Frontiers of two-dimensional correlation spectroscopy. Part 1. New concepts and noteworthy developments. *J. Mol. Struct.* 1069 (2014) 3, <http://dx.doi.org/10.1016/j.molstruc.2014.01.025>.
- [4] I. Noda, Frontiers of two-dimensional correlation spectroscopy. Part 2. Perturbation methods, fields of applications, and types of analytical probes. *J. Mol. Struct.* 1069 (2014) 23, <http://dx.doi.org/10.1016/j.molstruc.2014.01.016>.

- [5] L.J. Frasinski, K. Codling, P.A. Hatherly, Covariance mapping: a correlation method applied to multiphoton multiple ionization, *Science* 246 (1989) 1029.
- [6] I. Noda, Two-dimensional infrared spectroscopy, *J. Am. Chem. Soc.* 111 (1989) 8116.
- [7] I. Noda, Two-dimensional infrared (2D IR) spectroscopy: theory and applications, *Appl. Spectrosc.* 44 (1990) 550.
- [8] I. Noda, Generalized two-dimensional correlation method applicable to infrared, Raman, and other types of spectroscopy, *Appl. Spectrosc.* 47 (1993) 1329.
- [9] R.R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford, 1992.
- [10] A. Bax, *Two-Dimensional Nuclear Magnetic Resonance in Liquids*, D. Reidel Publishing Company, Dordrecht, Holland, 1982.
- [11] R. Bruschweiler, Theory of covariance nuclear magnetic resonance spectroscopy, *J. Chem. Phys.* 121 (2004) 409.
- [12] I. Noda, A.E. Dowrey, C. Marcott, G.M. Story, Y. Ozaki, Generalized two-dimensional correlation spectroscopy, *Appl. Spectrosc.* 54 (2000) 236A.
- [13] F.-L. Zhang, L. Bruschweiler-Li, R. Bruschweiler, Simultaneous de novo identification of molecules in chemical mixtures by doubly indirect covariance NMR spectroscopy, *J. Am. Chem. Soc.* 132 (2010) 16922.
- [14] D.A. Snyder, R. Bruschweiler, Generalized indirect covariance NMR formalism for establishment of multidimensional spin correlations, *J. Phys. Chem. A* 113 (2009) 12898.
- [15] K. Bingol, R.K. Salinas, R. Brueschweiler, Higher-rank correlation NMR spectra with spectral moment filtering, *J. Phys. Chem. Lett.* 1 (2010) 1086.
- [16] D.A. Snyder, F. Zhang, R. Brueschweiler, Covariance NMR in higher dimensions: application to 4D NOESY spectroscopy of proteins, *J. Biomol. NMR* 39 (2007) 165.
- [17] Y. Chen, F. Zhang, D. Snyder, Z. Gan, L. Bruschweiler-Li, R. Brueschweiler, Quantitative covariance NMR by regularization, *J. Biomol. NMR* 38 (2007) 73.
- [18] K.A. Blinov, N.I. Larin, M.P. Kvasha, A. Moser, A.J. Williams, G.E. Martin, Analysis and elimination of artifacts in indirect covariance NMR spectra via unsymmetrical processing, *Magn. Reson. Chem.* 43 (2005) 999.
- [19] K.A. Blinov, N.I. Larin, A.J. Williams, K.A. Mills, G.E. Martin, Unsymmetrical covariance processing of COSY or TOCSY and HSQC NMR data to obtain the equivalent of HSQC-COSY or HSQC-TOCSY spectra, *J. Heterocyclic Chem.* 43 (2006) 163.
- [20] G.E. Martin, B.D. Hilton, P.A. Irish, K.A. Blinov, A.J. Williams, Using unsymmetrical indirect covariance processing to calculate GHSQC-COSY spectra, *J. Nat. Prod.* 70 (2007) 1393.
- [21] F. Zhang, R. Brueschweiler, Indirect covariance NMR spectroscopy, *J. Am. Chem. Soc.* 126 (2004) 13180.
- [22] K.A. Blinov, N.I. Larin, A.J. Williams, M. Zell, G.E. Martin, Long-range carbon–carbon connectivity via unsymmetrical indirect covariance processing of HSQC and HMBC NMR data, *Magn. Reson. Chem.* 44 (2006) 107.
- [23] K.A. Blinov, A.J. Williams, B.D. Hilton, P.A. Irish, G.E. Martin, The use of unsymmetrical indirect covariance NMR methods to obtain the equivalent of HSQC-NOESY data, *Magn. Reson. Chem.* 45 (2007) 544.
- [24] R.L.E.G. Aspers, P.E.T.J. Geutjes, M. Honing, M. Jaeger, Using indirect covariance processing for structure elucidation of small molecules in cases of spectral crowding, *Magn. Reson. Chem.* 49 (2011) 425.
- [25] G.E. Martin, B.D. Hilton, K.A. Blinov, A.J. Williams, Unsymmetrical indirect covariance processing of hyphenated and long-range heteronuclear 2D NMR spectra-enhanced visualization of  $^2J_{\text{CH}}$  and  $^4J_{\text{CH}}$  correlation responses, *J. Heterocyclic Chem.* 45 (2008) 1109.

- [26] E.R. Zartler, G.E. Martin, The use of  $^1\text{H}$ - $^{31}\text{P}$  GHMBC and covariance NMR to unambiguously determine phosphate ester linkages in complex polysaccharide mixtures, *J. Biomol. NMR* 51 (2011) 357.
- [27] G.E. Martin, B.D. Hilton, K.A. Blinov, A.J. Williams,  $^{13}\text{C}$ - $^{15}\text{N}$  correlation via unsymmetrical indirect covariance NMR: application to vinblastine, *J. Nat. Prod.* 70 (2007) 1966.
- [28] G.E. Martin, B.D. Hilton, K.A. Blinov, HSQC-ADEQUATE correlation: a new paradigm for establishing a molecular skeleton, *Magn. Reson. Chem.* 49 (2011) 248.
- [29] G.E. Martin, B.D. Hilton, R.M. Willcott III, K.A. Blinov, HSQC-1, n-ADEQUATE: a new approach to long-range  $^{13}\text{C}$ - $^{13}\text{C}$  correlation by covariance processing, *Magn. Reson. Chem.* 49 (2011) 641.
- [30] G.E. Martin, W.R. Thomas, K.A. Blinov, C.G. Anklin, W. Bermel, HMBC-1,1-ADEQUATE via generalized indirect covariance: a high sensitivity alternative to n,1-ADEQUATE, *Magn. Reson. Chem.* 50 (2012) 691.
- [31] G.E. Martin, K.A. Blinov, M. Reibarkh, R.T. Williamson,  $^1\text{J}_{\text{CC}}$ -edited HSQC-1, n-ADEQUATE: a new paradigm for simultaneous direct and long-range carbon-carbon correlation, *Magn. Reson. Chem.* 50 (2012) 722.
- [32] B. Hu, J.-P. Amoureaux, J. Trebosc, M. Deschamps, G. Tricot, Solid-state NMR covariance of homonuclear correlation spectra, *J. Chem. Phys.* 128 (2008) 134502/1.
- [33] G.E. Martin, B.D. Hilton, K.A. Blinov, HSQC-1,1-ADEQUATE and HSQC-1, n-ADEQUATE: enhanced methods for establishing adjacent and long-range  $^{13}\text{C}$ - $^{13}\text{C}$  connectivity networks, *J. Nat. Prod.* 74 (2011) 2400.
- [34] K. Takeda, Y. Kusakabe, Y. Noda, M. Fukuchi, K. Takegoshi, Homo- and heteronuclear two-dimensional covariance solid-state NMR spectroscopy with a dual-receiver system, *Phys. Chem. Chem. Phys.* 14 (2012) 9715.
- [35] N. Trbovic, S. Smirnov, F. Zhang, R. Brueschweiler, Covariance NMR spectroscopy by singular value decomposition, *J. Magn. Reson.* 171 (2004) 277.
- [36] D.A. Snyder, A. Ghosh, F. Zhang, T. Szyperski, R. Bruschweiler, Z-matrix formalism for quantitative noise assessment of covariance nuclear magnetic resonance spectra, *J. Chem. Phys.* 129 (2008) 104511.
- [37] R. Bruschweiler, F. Zhang, Covariance nuclear magnetic resonance spectroscopy, *J. Chem. Phys.* 120 (2004) 5253.
- [38] O. Lafon, B. Hu, J.-P. Amoureaux, P. Lesot, Fast and high-resolution stereochemical analysis by nonuniform sampling and covariance processing of anisotropic natural abundance 2D 2H NMR datasets, *Chem. Eur. J.* 17 (2011) 6716.
- [39] B.-W. Hu, P. Zhou, I. Noda, G.-Z. Zhao, An NMR approach applicable to biomolecular structure characterization, *Anal. Chem.* 77 (2005) 7534.
- [40] G.E. Martin, B.D. Hilton, P.A. Irish, K.A. Blinov, A.J. Williams,  $^{13}\text{C}$ - $^{15}\text{N}$  connectivity networks via unsymmetrical indirect covariance processing of  $^1\text{H}$ - $^{13}\text{C}$  HSQC and  $^1\text{H}$ - $^{15}\text{N}$  IMPEACH spectra, *J. Heterocyclic Chem.* 44 (2007) 1219.
- [41] G.E. Martin, B.D. Hilton, P.A. Irish, K.A. Blinov, A.J. Williams, Application of unsymmetrical indirect covariance NMR methods to the computation of the  $^{13}\text{C} \leftrightarrow ^{15}\text{N}$  HMBC-IMPEACH correlation spectra, *Magn. Reson. Chem.* 45 (2007) 883.
- [42] G.E. Martin, P.A. Irish, B.D. Hilton, K.A. Blinov, A.J. Williams, Utilizing unsymmetrical indirect covariance processing to define  $^{15}\text{N}$ - $^{13}\text{C}$  connectivity networks, *Magn. Reson. Chem.* 45 (2007) 624.
- [43] G.E. Martin, B.D. Hilton, K.A. Blinov, A.J. Williams, Using indirect covariance spectra to identify artifact responses in unsymmetrical indirect covariance calculated spectra, *Magn. Reson. Chem.* 46 (2008) 138.
- [44] G.M. Kirwan, M.J. Adams, Peak width issues with generalised 2D correlation NMR spectroscopy, *J. Mol. Struct.* 892 (2008) 225.

- [45] G.E. Martin, B.D. Hilton, K.A. Blinov, A.J. Williams, Multistep correlations via covariance processing of COSY/GCOSY spectra: opportunities and artifacts, *Magn. Reson. Chem.* 46 (2008) 997.
- [46] S. Sasic, Two-dimensional correlation analysis of nuclear magnetic resonance metabonomics data, *Appl. Spectrosc.* 62 (2008) 840.
- [47] Y. Chen, F. Zhang, R. Bruschweiler, Residual water suppression by indirect covariance NMR, *Magn. Reson. Chem.* 45 (2007) 925.
- [48] I. Noda, Two-dimensional correlation spectroscopy—biannual survey 2007–2009, *J. Mol. Struct.* 974 (2010) 3.
- [49] I. Noda, G.M. Story, C. Marcott, Pressure-induced transitions of polyethylene studied by two-dimensional infrared correlation spectroscopy, *Vib. Spectrosc.* 19 (1999) 461.
- [50] C.D. Eads, I. Noda, Generalized correlation NMR spectroscopy, *J. Am. Chem. Soc.* 124 (2002) 1111.
- [51] H. Garcia, A.S. Barros, C. Gonçalves, F.M. Gama, A.M. Gil, Characterization of dextrin hydrogels by FTIR spectroscopy and solid state NMR spectroscopy, *Eur. Polym. J.* 44 (2008) 2318.
- [52] D.J. Crockford, E. Holmes, J.C. Lindon, R.S. Plumb, S. Zirah, S.J. Bruce, et al., Statistical heterospectroscopy, an approach to the integrated analysis of NMR and UPLC-MS data sets: application in metabonomic toxicology studies, *Anal. Chem.* 78 (2006) 363.
- [53] D.J. Crockford, A.D. Maher, K.R. Ahmadi, A. Barrett, R.S. Plumb, I.D. Wilson, et al., <sup>1</sup>H NMR and UPLC-MSE statistical heterospectroscopy: characterization of drug metabolites (xenometabolome) in epidemiological studies, *Anal. Chem.* 80 (2008) 6835.
- [54] C. Cobas, NMR data evaluation: review of covariance applications, *Stan's Libr.* 1 (2014).
- [55] S.L. Robinette, F. Zhang, L. Bruschweiler-Li, R. Bruschweiler, Web server based complex mixture analysis by NMR, *Anal. Chem.* 80 (2008) 3606.
- [56] F. Zhang, S.L. Robinette, L. Bruschweiler-Li, R. Brueschweiler, Web server suite for complex mixture analysis by covariance NMR, *Magn. Reson. Chem.* 47 (2009) S118.
- [57] T. Short, L. Alzapiedi, R. Bruschweiler, D. Snyder, A covariance NMR toolbox for MATLAB and OCTAVE, *J. Magn. Reson.* 209 (2011) 75.
- [58] J. Helmus, C. Jaroniec, Nmrglue: an open source Python package for the analysis of multidimensional NMR data, *J. Biomol. NMR* 55 (2013) 355.
- [59] F. Delaglio, S. Grzesiek, G. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multi-dimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6 (1995) 277.
- [60] R. Bruschweiler, F. Zhang, Robust deconvolution of complex mixtures by covariance spectroscopy, *Angew. Chem. Int. Ed.* 46 (2007) 2639–2642.
- [61] K. Bingol, F. Zhang, L. Bruschweiler-Li, R. Bruschweiler, TOCCATA: a customized carbon total correlation spectroscopy NMR metabolomics database, *Anal. Chem.* 84 (2012) 9395.
- [62] I. Noda, Double two-dimensional correlation analysis—2D correlation of 2D spectra, *J. Mol. Struct.* 974 (2010) 108.
- [63] S. Hyberts, H. Arthanari, G. Wagner, Applications of non-uniform sampling and processing, in: M. Billeter, V. Orekhov (Eds.), *Nov. Sampl. Approaches High. Dimens. NMR SE-187*, Springer, Berlin, Heidelberg, 2012, pp. 125–148.
- [64] P. Schmieder, A. Stern, G. Wagner, J. Hoch, Application of nonlinear sampling schemes to COSY-type spectra, *J. Biomol. NMR* 3 (1993) 569.
- [65] M. Maciejewski, M. Mobli, A. Schuyler, A. Stern, J. Hoch, Data sampling in multi-dimensional NMR: fundamentals and strategies, in: M. Billeter, V. Orekhov (Eds.),

- Nov. Sampl. Approaches High. Dimens. NMR SE-185, Springer, Berlin, Heidelberg, 2012, pp. 49–77.
- [66] K. Kazimierczuk, M. Misiak, J. Stanek, A. Zawadzka-Kazimierczuk, W. Koźmiński, Generalized Fourier transform for non-uniform sampled data, in: M. Billeter, V. Orekhov (Eds.), Nov. Sampl. Approaches High. Dimens. NMR SE-186, Springer, Berlin, Heidelberg, 2012, pp. 79–124.
- [67] J.C. Hoch, M.W. Maciejewski, M. Mobli, A.D. Schuyler, A.S. Stern, Nonuniform sampling and maximum entropy reconstruction in multidimensional NMR, *Acc. Chem. Res.* 47 (2014) 708.
- [68] Y. Chen, F. Zhang, W. Bermel, R. Brueschweiler, Enhanced covariance spectroscopy from minimal datasets, *J. Am. Chem. Soc.* 128 (2006) 15564.
- [69] D. Merlet, B. Ancian, J. Courtieu, P. Lesot, Two-dimensional deuterium NMR spectroscopy of chiral molecules oriented in a polypeptide liquid crystal: applications for the enantiomeric analysis through natural abundance deuterium NMR, *J. Am. Chem. Soc.* 121 (1999) 5249.
- [70] O. Lafon, P. Lesot, D. Merlet, J. Courtieu, Modified z-gradient filtering as a mean to obtain phased deuterium autocorrelation 2D NMR spectra in oriented solvents, *J. Magn. Reson.* 171 (2004) 135.
- [71] J.C.J. Barna, E.D. Laue, M.R. Mayger, J. Skilling, S.J.P. Worrall, Exponential sampling, an alternative method for sampling in two-dimensional NMR experiments, *J. Magn. Reson.* 73 (1987) 69.
- [72] Y. Li, Q. Wang, Z. Zhang, J. Yang, B. Hu, Q. Chen, et al., Covariance spectroscopy with a non-uniform and consecutive acquisition scheme for signal enhancement of the NMR experiments, *J. Magn. Reson.* 217 (2012) 106.
- [73] W. Qiang, Signal enhancement for the sensitivity-limited solid state NMR experiments using a continuous, non-uniform acquisition scheme, *J. Magn. Reson.* 213 (2011) 171.
- [74] A. Kumar, S.C. Brown, M.E. Donlan, B.U. Meier, P.W. Jeffs, Optimization of two-dimensional NMR by matched accumulation, *J. Magn. Reson.* 95 (1991) 1.
- [75] D. Rovnyak, M. Sarcone, Z. Jiang, Sensitivity enhancement for maximally resolved two-dimensional NMR by nonuniform sampling, *Magn. Reson. Chem.* 49 (2011) 483.
- [76] Y. Li, B. Hu, Q. Chen, Q. Wang, Z. Zhang, J. Yang, et al., Comparison of various sampling schemes and accumulation profiles in covariance spectroscopy with exponentially decaying 2D signals, *Analyst* 138 (2013) 2411.
- [77] E.C. Lin, S.J. Opella, Covariance spectroscopy in high-resolution multi-dimensional solid-state NMR, *J. Magn. Reson.* 239 (2014) 57.
- [78] S.H. Park, S. Berkamp, G.A. Cook, M.K. Chan, H. Viadiu, S.J. Opella, Nanodiscs versus macrodiscs for NMR of membrane proteins, *Biochemistry* 50 (2011) 8983.
- [79] K. Kazimierczuk, V.Y. Orekhov, Accelerated NMR spectroscopy by using compressed sensing, *Angew. Chem. Int. Ed.* 50 (2011) 5556.
- [80] F. Zhang, N. Trbovic, J. Wang, R. Brueschweiler, Double-quantum biased covariance spectroscopy: application to the 2D INADEQUATE experiment, *J. Magn. Reson.* 174 (2005) 219.
- [81] F. Zhang, L. Bruschweiler-Li, R. Brueschweiler, High-resolution homonuclear 2D NMR of carbon-13 enriched metabolites and their mixtures, *J. Magn. Reson.* 225 (2012) 10.
- [82] F. Zhang, R. Brueschweiler, Spectral deconvolution of chemical mixtures by covariance NMR, *ChemPhysChem* 5 (2004) 794.
- [83] F. Zhang, R. Brueschweiler, Robust deconvolution of complex mixtures by covariance TOCSY spectroscopy, *Angew. Chem. Int. Ed.* 46 (2007) 2639.

- [84] F. Zhang, A.T. Dossey, C. Zachariah, A.S. Edison, R. Brueschweiler, Strategy for automated analysis of dynamic metabolic mixtures by NMR. Application to an insect venom, *Anal. Chem.* 79 (2007) 7748.
- [85] F. Zhang, L. Bruschweiler-Li, S.L. Robinette, R. Bruschweiler, Self-consistent metabolic mixture analysis by heteronuclear NMR. application to a human cancer cell line, *Anal. Chem.* 80 (2008) 7549.
- [86] M. Weingarth, P. Tekely, R. Brueschweiler, G. Bodenhausen, Improving the quality of 2D solid-state NMR spectra of microcrystalline proteins by covariance analysis, *Chem. Commun.* 46 (2010) 952.
- [87] Y.-I. Masuda, M. Fukuchi, T. Yatagawa, M. Tada, K. Takeda, K. Irie, et al., Solid-state NMR analysis of interaction sites of curcumin and 42-residue amyloid  $\beta$ -protein fibrils, *Bioorg. Med. Chem.* 19 (2011) 5967.
- [88] W. Schoefberger, V. Smrecki, D. Vikic-Topic, N. Mueller, Homonuclear long-range correlation spectra from HMBC experiments by covariance processing, *Magn. Reson. Chem.* 45 (2007) 583.
- [89] G.A. Morris, J.A. Aguilar, R. Evans, S. Haiber, M. Nilsson, True chemical shift correlation maps: a TOCSY experiment with pure shifts in both dimensions, *J. Am. Chem. Soc.* 132 (2010) 12770.
- [90] K. Zanger, H. Sterk, Homonuclear broadband-decoupled NMR spectra, *J. Magn. Reson.* 124 (1997) 486.
- [91] J.A. Aguilar, A.A. Colbourne, J. Cassani, M. Nilsson, G.A. Morris, Decoupling two-dimensional NMR spectroscopy in both dimensions: pure shift NOESY and COSY, *Angew. Chem. Int. Ed.* 51 (2012) 6460.
- [92] B. Hu, J.-P. Amoureaux, J. Trebosc, Indirect covariance NMR spectroscopy of through-bond homo-nuclear correlations for quadrupolar nuclei in solids under high-resolution, *Solid State Nucl. Magn. Reson.* 31 (2007) 163.
- [93] G.E. Martin, Small-volume and high-sensitivity NMR probes, *Annu. Rep. NMR Spectrosc.* 56 (2005) 1.
- [94] B.D. Hilton, G.E. Martin, Investigation of the experimental limits of small-sample heteronuclear 2D NMR, *J. Nat. Prod.* 73 (2010) 1465.
- [95] G.E. Martin, B.D. Hilton, M.R. Willcott III, K.A. Blinov, HSQC-ADEQUATE: an investigation of data requirements, *Magn. Reson. Chem.* 49 (2011) 350.
- [96] G.E. Martin, D. Sunseri, Establishing the carbon skeleton of pharmaceutical agents using HSQC-ADEQUATE spectra, *J. Pharm. Biomed. Anal.* 55 (2011) 895.
- [97] G.E. Martin, Posaconazole: application of HSQC-ADEQUATE from general indirect covariance processing, *J. Heterocyclic Chem.* 49 (2012) 716.
- [98] W. Zhong, X. Yang, W. Tong, G.E. Martin, Structural characterization of a novel degradant of the antifungal agent posaconazole, *J. Pharm. Biomed. Anal.* 66 (2012) 40.
- [99] K. Bingol, R. Bruschweiler, Deconvolution of chemical mixtures with high complexity by NMR consensus trace clustering, *Anal. Chem.* 83 (2011) 7412.
- [100] D.A. Snyder, Y. Xu, D. Yang, R. Brueschweiler, Resolution-enhanced 4D  $^{15}\text{N}/^{13}\text{C}$  NOESY protein NMR spectroscopy by application of the covariance transform, *J. Am. Chem. Soc.* 129 (2007) 14126.
- [101] S. Pudakalakatti, A. Dubey, G. Jaipuria, U. Shubhashree, S. Adiga, D. Moskau, et al., A fast NMR method for resonance assignments: application to metabolomics, *J. Biomol. NMR* (2014) 1.
- [102] A. Chen, C.M. Brown, Distinct families of cis-acting RNA replication epsilon elements from hepatitis B viruses, *RNA Biol.* 9 (2012) 130.
- [103] G.-B. Luo, M. Karplus, Determining the conformational change that accompanies donor-acceptor distance fluctuations: an umbrella sampling analysis, *J. Phys. Chem. B* 115 (2011) 7991.

- [104] R. Selvaratnam, S. Chowdhury, S.B. Van, G. Melacini, Mapping allostery through the covariance analysis of NMR chemical shifts, *Proc. Natl. Acad. Sci. U.S.A.* (2011) 1.
- [105] G. Benison, D.S. Berkholz, E. Barbar, Protein assignments without peak lists using higher-order spectra, *J. Magn. Reson.* 189 (2007) 173.
- [106] D.J. Rigden, Use of covariance analysis for the prediction of structural domain boundaries from multiple protein sequence alignments, *Protein Eng.* 15 (2002) 65.
- [107] J.J. Prompers, R. Brueschweiler, General framework for studying the dynamics of folded and nonfolded proteins by NMR relaxation spectroscopy and MD simulation, *J. Am. Chem. Soc.* 124 (2002) 4522.
- [108] I. Noda, Close-up view on the inner workings of two-dimensional correlation spectroscopy, *Vib. Spectrosc.* 60 (2012) 146.
- [109] I. Noda, Recent advancement in the field of two-dimensional correlation spectroscopy, *J. Mol. Struct.* 883–884 (2008) 2.
- [110] M. Edgar, Physical methods and techniques: NMR spectroscopy, *Annu. Rep. B Org. Chem.* 109 (2013) 256.
- [111] M. Edgar, Physical methods and techniques NMR spectroscopy, *Annu. Rep. B Org. Chem.* 108 (2012) 292.
- [112] M. Edgar, Physical methods and techniques: NMR spectroscopy, *Annu. Rep. B Org. Chem.* 107 (2011) 308.
- [113] M. Edgar, Physical methods and techniques NMR spectroscopy, *Annu. Rep. B Org. Chem.* 106 (2010) 325.
- [114] M. Edgar, Physical methods and techniques NMR spectroscopy, *Annu. Rep. B Org. Chem.* 105 (2009) 340.
- [115] M. Edgar, Physical methods and techniques: NMR spectroscopy, *Annu. Rep. Prog. Chem. B Org. Chem.* 104 (2008) 312.
- [116] G.M. Kirwan, S. Clark, N.W. Barnett, J.O. Niere, M.J. Adams, Generalised 2D-correlation NMR analysis of a wine fermentation, *Anal. Chim. Acta* 629 (2008) 128.
- [117] J.C. Lindon, J.K. Nicholson, E. Holmes, *The Handbook of Metabonomics and Metabolomics*, Elsevier B.V., Amsterdam, Oxford, 2007
- [118] S.L. Robinette, J.C. Lindon, J.K. Nicholson, Statistical spectroscopic tools for biomarker discovery and systems medicine, *Anal. Chem.* 85 (2013) 5297.
- [119] S.L. Robinette, R. Bruschweiler, F.C. Schroeder, A.S. Edison, NMR in metabolomics and natural products research: two sides of the same coin, *Acc. Chem. Res.* 45 (2012) 288.
- [120] O. Cloarec, M.-E. Dumas, A. Craig, R.H. Barton, J. Trygg, J. Hudson, et al., Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic  $^1\text{H}$  NMR data sets, *Anal. Chem.* 77 (2005) 1282.
- [121] E. Holmes, O. Cloarec, J.K. Nicholson, Probing latent biomarker signatures and in vivo pathway activity in experimental disease states via statistical total correlation spectroscopy (STOCSY) of biofluids: application to  $\text{HgCl}_2$  toxicity, *J. Proteome Res.* 5 (2006) 1313.
- [122] J.M. Posma, I. Garcia-Perez, M. De Iorio, J.C. Lindon, P. Elliott, E. Holmes, et al., Subset optimization by reference matching (STORM): an optimized statistical approach for recovery of metabolic biomarker structural information from  $^1\text{H}$  NMR spectra of biofluids, *Anal. Chem.* 84 (2012) 10694.
- [123] L.M. Smith, A.D. Maher, O. Cloarec, M. Rantalainen, H. Tang, P. Elliott, et al., Statistical correlation and projection methods for improved information recovery from diffusion-edited NMR spectra of biological samples, *Anal. Chem.* 79 (2007) 5682.
- [124] J.M. Fonville, A.D. Maher, M. Coen, E. Holmes, J.C. Lindon, J.K. Nicholson, Evaluation of full-resolution  $J$ -resolved  $^1\text{H}$  NMR projections of biofluids for

- metabonomics information retrieval and biomarker identification, *Anal. Chem.* 82 (2010) 1811.
- [125] M. Coen, Y.-S. Hong, O. Cloarec, C.M. Rhode, M.D. Reily, D.G. Robertson, et al., Heteronuclear  $^1\text{H}$ - $^{31}\text{P}$  statistical total correlation NMR spectroscopy of intact liver for metabolic biomarker assignment: application to galactosamine-induced hepatotoxicity, *Anal. Chem.* 79 (2007) 8956.
- [126] H.C. Keun, T.J. Athersuch, O. Beckonert, Y. Wang, J. Saric, J.P. Shockcor, et al., Heteronuclear  $^{19}\text{F}$ - $^1\text{H}$  statistical total correlation spectroscopy as a tool in drug metabolism: study of flucloxacillin biotransformation, *Anal. Chem.* 80 (2008) 1073.
- [127] S.L. Robinette, K.A. Veselkov, E. Bohus, M. Coen, H.C. Keun, T.M.D. Ebbels, et al., Cluster analysis statistical spectroscopy using nuclear magnetic resonance generated metabolic data sets from perturbed biological systems, *Anal. Chem.* 81 (2009) 6581.
- [128] C.H. Johnson, T.J. Athersuch, I.D. Wilson, L. Iddon, X. Meng, A.V. Stachulski, et al., Kinetic and J-resolved statistical total correlation NMR spectroscopy approaches to structural information recovery in complex reacting mixtures: application to acyl glucuronide intramolecular transacylation reactions, *Anal. Chem.* 80 (2008) 4886.
- [129] C.J. Sands, M. Coen, T.M.D. Ebbels, E. Holmes, J.C. Lindon, J.K. Nicholson, Data-driven approach for metabolite relationship recovery in biological  $^1\text{H}$  NMR data sets using iterative statistical total correlation spectroscopy, *Anal. Chem.* 83 (2011) 2075.
- [130] B.J. Blaise, V. Navratil, C. Domange, L. Shintu, M.-E. Dumas, B. Elena-Herrmann, et al., Two-dimensional statistical recoupling for the identification of perturbed metabolic networks from NMR spectroscopy, *J. Proteome Res.* 9 (2010) 4513.
- [131] B.J. Blaise, V. Navratil, L. Emsley, P. Toulhoat, Orthogonal filtered recoupled-STOCSY to extract metabolic networks associated with minor perturbations from NMR spectroscopy, *J. Proteome Res.* 10 (2011) 4342.
- [132] C.J. Sands, M. Coen, A.D. Maher, T.M.D. Ebbels, E. Holmes, J.C. Lindon, et al., Statistical total correlation spectroscopy editing of  $^1\text{H}$  NMR spectra of biofluids: application to drug metabolite profile identification and enhanced information recovery, *Anal. Chem.* 81 (2009) 6458.
- [133] A.D. Maher, J.M. Fonville, M. Coen, J.C. Lindon, C.D. Rae, J.K. Nicholson, Statistical total correlation spectroscopy scaling for enhancement of metabolic information recovery in biological NMR spectra, *Anal. Chem.* 84 (2011) 1083.
- [134] B.J. Blaise, L. Shintu, B. Elena, L. Emsley, M.-E. Dumas, P. Toulhoat, Statistical recoupling prior to significance testing in nuclear magnetic resonance based metabonomics, *Anal. Chem.* 81 (2009) 6242.
- [135] J.C. Lindon, J.K. Nicholson, Spectroscopic and statistical techniques for information recovery in metabonomics and metabolomics, *Annu. Rev. Anal. Chem.* 1 (2008) 45.
- [136] E.M. Lenz, I.D. Wilson, Analytical strategies in metabonomics, *J. Proteome Res.* 6 (2007) 443.
- [137] R. Huo, de M.R.A. van, J.A. Pikkemaat, R. Wehrens, L.M.C. Buydens, Diagnostic analysis of experimental artifacts in DOSY NMR data by covariance matrix of the residuals, *J. Magn. Reson.* 172 (2005) 346.
- [138] A.A. Colbourne, G.A. Morris, M. Nilsson, Local covariance order diffusion-ordered spectroscopy: a powerful tool for mixture analysis, *J. Am. Chem. Soc.* 133 (2011) 7640.
- [139] D.L.A.G. Griminck, S.K. Vasa, W.L. Meerts, A.P.M. Kentgens, A. Brinkmann, EASY-GOING DUMBO on-spectrometer optimisation of phase modulated homonuclear decoupling sequences in solid-state NMR, *Chem. Phys. Lett.* 509 (2011) 186.
- [140] A.C.J. Weber, E.E. Burnell, Conformational statistics of *n*-butane in the condensed phase and the effects of temperature, *Chem. Phys. Lett.* 506 (2011) 196.

- [141] A.C.J. Weber, A. Pizzirusso, L. Muccioli, C. Zannoni, W.L. Meerts, C.A. De lange, et al., Efficient analysis of highly complex nuclear magnetic resonance spectra of flexible solutes in ordered liquids by using molecular dynamics, *J. Chem. Phys.* 136 (2012) 174506/1.
- [142] J. Fukazawa, K. Takegoshi, Phase covariance in NMR signal, *Phys. Chem. Chem. Phys.* 12 (2010) 11225.