J Math Chem (2013) 51:1961–1980
DOI 10.1007/s10910-013-0191-2

ORIGINAL PAPER

Aliasing functions and a dynamic approach to optimizing spectral widths in 2D experiments

José Félix Gómez-Reyes · Armando Ariza-Castolo

Received: 26 October 2012 / Accepted: 15 May 2013 / Published online: 4 June 2013

Abstract We reduce the modulo function that describes the aliasing of an NMR signal to a floor function form. An analysis of this function is carried out and an expression that defines its points of discontinuity is derived. Based only on this definition we develop a new method to optimize the spectral widths in the ^{13}C dimension of heteronuclear 2D experiments of small organic compounds using neither the modulo function nor the definition of the overlap points. We apply this method to the carbon chemical shifts of cholesterol, and find that it unambiguously assigns all of the signals acquired in the aliased Heteronuclear Single Quantum Coherence spectrum of cholesterol with the calculated spectral width. Previous reports do not show a fully resolved 2D spectrum of cholesterol.

Keywords Aliasing · Spectral widths optimization · High resolution in indirect dimensions · 2D-NMR spectroscopy

1 Introduction

Heteronuclear Single Quantum Coherence (HSQC) and Heteronuclear Multiple Bond Correlation (HMBC) pulse sequences are widely used in 2D NMR spectroscopy to

Electronic supplementary material The online version of this article (doi:10.1007/s10910-013-0191-2) contains supplementary material, which is available to authorized users.

J. F. Gómez-Reyes (✉) · A. Ariza-Castolo (✉)
Departamento de Química, Centro de Investigación y de Estudios Avanzados del IPN, Av. IPN 2508, Col. San Pedro Zacatenco, C.P. 07360 Mexico, DF, Mexico
e-mail: aariza@cinvestav.mx

J. F. Gómez-Reyes
e-mail: jfgomez@cinvestav.mx

assign chemical shifts (CSs) to organic molecules. A persistent problem, however, is a low resolution in the indirectly detected dimension. Reducing the spectral width is one way to increase this resolution. It is also possible to reduce the experimental time by simultaneously reducing the number of time increments and the observation range [5]. One disadvantage of reducing the spectral width is that resonances appear at frequencies different to their real CSs in the aliased spectra.

In a significant development for HSQC experiment, Jeannerat [12] proposed a method that greatly improved resolution in ^{13}C dimension by reducing the observed spectral width below the ^{13}C spectral width, thereby violating the Nyquist condition. This reduced signal crowding, but resulted in aliased spectra. Since then, algorithms have been developed to determine which of the aliased spectra generated with this method give the correct CSs. The early version of this method was the first to obtain a reference ^{13}C spectrum as well as a full reference HSQC spectrum, and choose a spectral window to be an arbitrary fraction of the width of the reference spectrum [12]. A later version, which required only one reference spectrum (^{13}C) to predict the best narrow spectral window [13], had the advantage of resolving spectra aliased by factors over 100.

The multiple-aliasing method was further improved [14,11] by using an algorithm, based on a list of ^{13}C CSs from a reference 1D spectrum, to calculate not only the optimal spectral width to use in a 2D acquisition, but also the optimal number of time increments in the carbon dimension. This version of the method achieved both higher resolution and a much lower experimental time because of the optimized time increments. In this algorithm, the location of each aliased signal is given by the modulo function (MF) using the binary operation mod [8] to relate the apparent frequency of a resonance to its real frequency:

$$\nu_a = mod\left(\nu_0 + \frac{SW_a}{2} - CF, SW_a\right) - \frac{SW_a}{2} + CF, \tag{1}$$

where SW_a is the spectral window considered, ν_0 the "the real" frequency and CF the carrier frequency in the F1 dimension. The spectral width of the overlap for any pair of signals i and j is predicted by the expression [14]

$$SW^* = \Delta^{ij}/n, \tag{2}$$

where n is an integer and $\Delta^{ij} = \nu_0^i - \nu_0^j$. These equations do not hold when the time proportional phase increment (TPPI) method is used to detect the signals [14]. Resonances that are not included in the observation range and are obtained with the TPPI method, appear as folded signals at the near border of the spectrum [3] instead of being wrapped around, which is the more common situation that we will review in this paper.

Jeannerat's algorithm [14] seeks to avoid signal overlaps. An overlap is first identified when the SW^* is within a range between high states of an overlap function. The tentative spectral width is reduced to correspond to the next lowest state of the overlap function, thereby resolving that particular pair. This is repeated for the other signals until either no more signals overlap or reducing the spectral width further

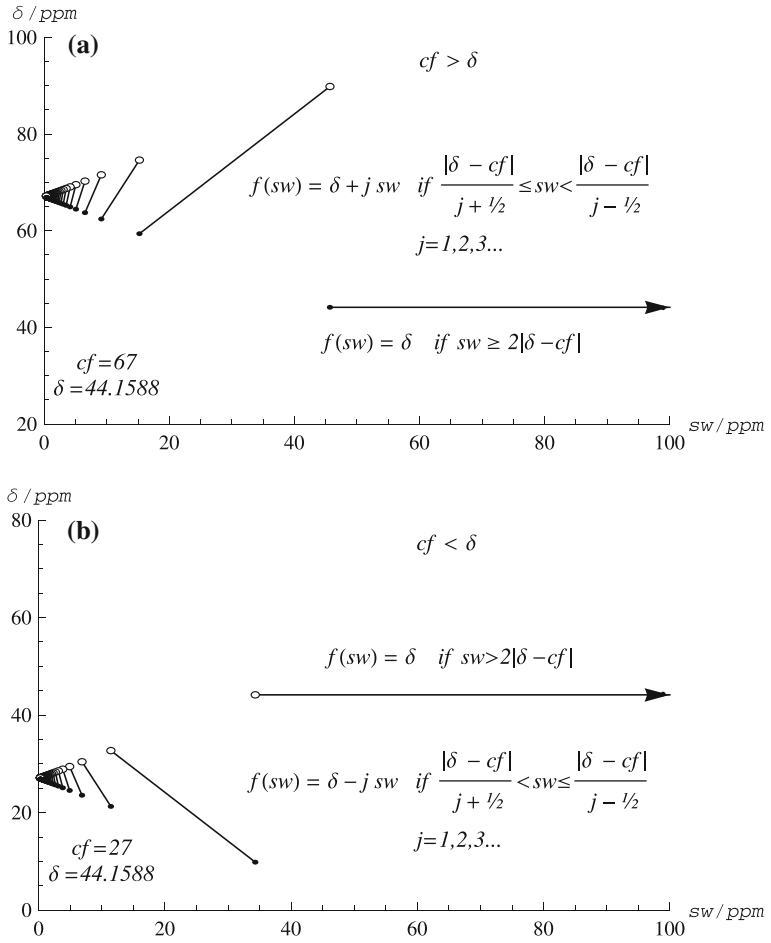

Fig. 1 Graph of the floor function (FF). The FF is reduced to δ when we set the values of cf and sw such that δ is inside sw. In the special cases where we set the values of cf and sw such that they are insufficient to include δ, we describe the output of the FF as a sequence of linear segments that differ in their slopes. The domains of the linear segments and the signs of their slopes depend on the position of cf with respect to δ. **a** Shows the case where $cf > \delta$, **b** shows the case where $cf < \delta$

cannot resolve any more signals. In the latter case, the number of time increments is increased and the entire process repeated. This continues until a fully resolved spectrum is obtained.

In Fig. 1, we present a sketch of the MF. The graph consists of an infinite set of linear segments, where the function "jumps" from one linear segment to another [7]. These particular points of abrupt change correspond to the points of discontinuity (PDs) of the function. The resonances at these frequencies appear at the borders of the spectrum, and depending on the resolution of the spectrum, parts of the signals may be missing. The definition of these PDs has not been previously described. In this study, we use a list of carbon CSs to define the PDs that affect each of the CSs. The expression that defines the PD of a particular CS also defines the slope that rules its

aliasing until the next PD in size. Calculating the complete set of PDs (from the set of CSs) and sorting them into decreasing numerical order, we can divide the domain of the MF into domain intervals (DIs). The slopes that rule the aliasing of the n CSs are constant in each of these DIs, and one can easily calculate the n slopes in each of these intervals. Proceeding this way, we calculate all that we need to know about the DIs, and then explore each DI by its set of linear equations instead of using the MF to search for a convenient spectral width free of overlaps and far from the borders of the spectrum.

The summary of the acronyms and variables definition that are used in this paper are given in "Appendix A and B", respectively. In this work, we have used structures known as lists [21,20] to collect the data (see "Appendix C"). The complete method is explained in detail with the aid of flowcharts. We provide a set of four user defined functions (UDFs), implemented in Mathematica® [21], that compute each stage of the method (Online Resource 1). UDFs, Build-in Mathematica® functions (BIMFs), signs and operators that are used in the flowchart expressions are given in "Appendix D".

2 Theory

2.1 Simplification of the MF

For the purposes of this study, we express Eq. (1) as

$$f(sw) = cf - \frac{sw}{2} + Mod\left[\delta - cf + \frac{sw}{2}, sw\right], \quad where \quad sw > 0, \tag{3}$$

where δ is the CS of a signal, cf the observation center and sw the observation range; all of these terms are expressed in parts per million (ppm). Defining the binary operation "mod" by means of the floor function [8], we obtain a reduced expression equivalent to Eq. (1) [7,6]:

$$f(sw) = cf - \frac{sw}{2} + \left(\delta - cf + \frac{sw}{2} - sw\left\lfloor \frac{\delta - cf + \frac{sw}{2}}{sw} \right\rfloor\right), \quad where\ sw > 0, \tag{4}$$

which is simplified to the Floor Function (FF):

$$f(sw) = \delta - sw\left\lfloor \frac{\delta - cf}{sw} + \frac{1}{2} \right\rfloor, \quad where \quad sw > 0. \tag{5}$$

2.2 Defining the domain of the FF by DIs

According to the value of δ, we can set the position and value of sw in two different ways when we perform an NMR experiment. Usually, we set sw and cf such that the peak of interest is included in sw:

$$|\delta - cf| < \frac{sw}{2}, \quad where \quad sw > 0. \tag{6}$$

Less typically, we set sw and cf in accordance with

$$|\delta - cf| \geq \frac{sw}{2}, where \quad cf \neq \delta. \tag{7}$$

In the first case, $|\delta - cf| < \frac{sw}{2}$ if and only if [17]

$$-\frac{sw}{2} < \delta - cf < \frac{sw}{2}. \tag{8}$$

Adding $\frac{sw}{2}$ to the inequality in Eq. (8), and multiplying the entire expression by sw^{-1}, we have

$$0 < \frac{\delta - cf}{sw} + \frac{1}{2} < 1. \tag{9}$$

Thus, considering Eq. (9), we obtain the following expression from the FF:

$$f(sw) = \delta - sw(0) = \delta. \tag{10}$$

We conclude that the FF is reduced to δ when we choose a value of sw that includes it.

In the second case, we can select cf and sw such that sw does not include the peak of interest. From (7), $|\delta - cf| \geq \frac{sw}{2}$ if and only if [17]

$$-(\delta - cf) = \frac{sw}{2} \quad or \quad -(\delta - cf) > \frac{sw}{2}, \quad where \quad cf > \delta, \tag{11}$$

$$or \quad \delta - cf \geq \frac{sw}{2}, \quad where \quad cf < \delta. \tag{12}$$

From the first part of (11), we have

$$0 = \frac{\delta - cf}{sw} + \frac{1}{2}, \quad where \quad cf > \delta. \tag{13}$$

Considering Eq. (13), we obtain the following from the FF:

$$f(sw) = \delta - sw(0) = \delta. \tag{14}$$

Therefore, the FF is reduced to δ when we choose sw according to the first part of Eq. (11). From the second part of 11, we obtain

$$\frac{\delta - cf}{sw} + \frac{1}{2} < 0, \quad where \quad cf > \delta. \tag{15}$$

The inequality in Eq. (15) can be written as

$$-j \leq \frac{\delta - cf}{sw} + \frac{1}{2} < -j + 1, \tag{16}$$

where $j = 1, 2, 3 \ldots$. Adding $-\frac{1}{2}$, taking the reciprocals of the terms and multiplying the entire expression by $(\delta - cf)$, we obtain the following inequality from (16):

$$\frac{|\delta - cf|}{j + \frac{1}{2}} \leq sw < \frac{|\delta - cf|}{j - \frac{1}{2}}. \tag{17}$$

From (16), if $-j \leq \frac{\delta - cf}{sw} + \frac{1}{2} < -j + 1$, then $\left\lfloor \frac{\delta - cf}{sw} + \frac{1}{2} \right\rfloor = -j$ [8] and we obtain the following equation from the FF:

$$f(sw) = \delta - sw(-j) = \delta + jsw. \tag{18}$$

Therefore, when $cf > \delta$ and sw varies in accordance with the second part of Eq. (11), the domain of the FF is divided into the intervals $\frac{|\delta - cf|}{j + \frac{1}{2}} \leq sw < \frac{|\delta - cf|}{j - \frac{1}{2}}$, where $j = 1, 2, 3 \ldots$. In each interval, the aliasing function has the form $\delta + j\,sw$. Figure 1a shows the graph of the FF for this case.

Similarly Eq. (12) transforms to the following inequalities:

$$j \leq \frac{\delta - cf}{sw} + \frac{1}{2} < j + 1 \tag{19}$$

and

$$\frac{|\delta - cf|}{j + \frac{1}{2}} < sw \leq \frac{|\delta - cf|}{j - \frac{1}{2}}. \tag{20}$$

From (19), if $j \leq \frac{\delta - cf}{sw} + \frac{1}{2} < j + 1$, then $\frac{\delta - cf}{sw} + \frac{1}{2} = j$, and we obtain the following equation from the FF:

$$f(sw) = \delta - sw(j) = \delta - j\,sw. \tag{21}$$

Thus, when $cf < \delta$ and sw varies in accordance with Eq. (12), the domain of the FF is split into the intervals $\frac{|\delta - cf|}{j + \frac{1}{2}} < sw \leq \frac{|\delta - cf|}{j - \frac{1}{2}}$, where $j = 1, 2, 3 \ldots$. In each interval, the FF is of the form $\delta - j\,sw$. Figure 1b shows the graph of the FF for this case.

From Eqs. (17), (18), (20) and (21), we ensure that between two consecutive PDs, the slope is constant. The OPs are not related to changes in the slopes. The set of PDs of the FF can be expressed by:

$$\text{set of PDs} = \left\{ \frac{|\delta - cf|}{j - 1/2} \right\}_{j=1}^{\infty}. \tag{22}$$

When $cf > \delta$, the *jth* PD is the left endpoint of the interval expressed by Eq. (17). When $cf < \delta$, the *jth* PD is the right endpoint of the interval expressed by Eq. (20). Whichever case we consider, where $cf > \delta$ or $cf < \delta$, the FF in each interval has the form:

$$\delta + (Sign\,[cf - \delta]\,j)\;\; sw \quad where \quad Sign\,[cf - \delta] = \begin{cases} -1 & if \quad cf < \delta \\ +1 & if \quad cf > \delta \end{cases}. \tag{23}$$

3 Calculation methods

3.1 Definition of the PDs

The values of the PDs define the endpoints of the DIs where we can search convenient spectral widths. To simplify calculation and description, it is helpful to first define the list of CSs values and the list of the PDs. The list of CSs is defined by:

$$\boldsymbol{\Delta} = \{\delta_i\}_{i=1}^{n}, \tag{24}$$

where n is the number of CS values listed in decreasing numerical order. From Eq. (22) we construct a convenient expression for the PDs for the list of CSs in Eq. (24) as follows:

$$\boldsymbol{E} = \left\{ \left\{ \frac{|\delta_i - cf|}{j - \frac{1}{2}}, i, j\ \mathrm{Sign}\,[cf - \delta_i] \right\}_{j=1}^{p_i} \right\}_{i=1}^{n}, \tag{25}$$

where

$$p_i = \frac{|\delta_i - cf|}{lmt} + \frac{1}{2}, |\delta_i - cf| \geq \frac{lmt}{2}. \tag{26}$$

The first part of each element is the size of the PD and represents one of the PDs in the graph of the FF of the ith CS value whose aliasing is ruled, until the next PD in size, by the slope encoded in the third part.

3.2 Definition of the DIs

The set of all PDs (Eq. 25) will be part of the list that defines the set of DIs, the list $\boldsymbol{\Phi}$:

$$\boldsymbol{\Phi} = \left\{ \left\{ \{\phi_{i11}, \phi_{i12}, \phi_{i13}\}, \{\phi_{i2k}\}_{k=1}^{n}, \{\phi_{i31}, \phi_{i32}\} \right\} \right\}_{i=1}^{p}, \tag{27}$$

where p is the number of elements in the list, i.e., the number of DIs in decreasing numerical order of their respective PDs. This expression contains all the information that we need to know about the DIs. The first part of each element of the list corresponds to one element ε_{ij} of list $\boldsymbol{E}$, from Eq. (25). Each second part stores the list of the n slopes that are subsequently calculated. The terms ϕ_{i31} and ϕ_{i32} are the size of the next PD (the term $\phi_{(i+1)11}$) and the difference $\phi_{i11} - \phi_{(i+1)11}$, respectively (in other words, the size of the current DI).

Figure 2 shows the flow chart of the procedure used to calculate the list $\boldsymbol{\Phi}$. First (Fig. 2a), the list of the differences between the observation center and each peak is computed and assigned to list $\boldsymbol{T}$. The list $\boldsymbol{T}$ is defined as a list of three values where the

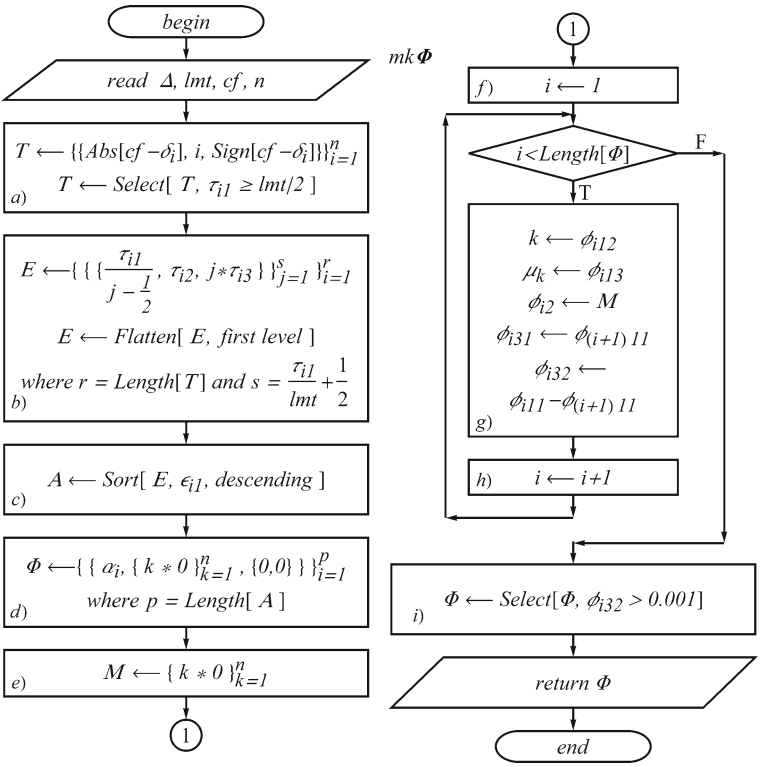

Fig. 2 Flowchart showing how to calculate the list of the DIs (***mkΦ***). First, the list of the absolute values of the differences between cf and each δ_i is assigned to the list $\boldsymbol{T}$ as a list of three values (**a**). We select the elements of the list $\boldsymbol{T}$ according to the *sw* limit. Then, we use the list $\boldsymbol{T}$ to calculate the list of the PDs (list $\boldsymbol{E}$ in **b**). The list $\boldsymbol{E}$ is used to make the list A, where the elements ε_{ij} are arranged in descending order by comparing their PD sizes (**c**). Once the list $\boldsymbol{A}$ is calculated, it is used to construct the list of the DIs (list $\boldsymbol{\Phi}$ in **d**). Next, we create the list $\boldsymbol{M}$ (**e**) that will be used for storing the slopes. The method used to calculate the set of slopes of each DI is represented by the loop in (**f–h**). Finally, we select the DIs according to their sizes (**i**). This method is described in more detail in the Sects. 3.1, 3.2 and 3.3

second elements encode the positions of their respective carbons on the list $\boldsymbol{\Delta}$, the "*i*th" value. The third elements will be used subsequently to calculate the slopes (see Eq. 23). The differences that are greater than or equal to half the spectral width limit are selected from $\boldsymbol{T}$. Second, from these elements we construct the list $\boldsymbol{E}$ (Fig. 2b). Third, we reorganize the elements ε_{ij} into the list $\boldsymbol{A}$ (Fig. 2c), where these lists are arranged in descending order by comparing their PD sizes. Next, the list $\boldsymbol{\Phi}$ is constructed based on the list $\boldsymbol{A}$ (Fig. 2d) and we calculate the set of slopes of each DI (Fig. 2f–h). Finally, we select the DIs (Fig. 2i). We do not explore DIs ≤ 0.001 ppm.

3.3 Calculating the slopes of the DIs

The set of slopes of the *i*th DI are stored in list $\boldsymbol{M}$, which is set to zero at the beginning of the process (Fig. 2e). Following list $\boldsymbol{\Phi}$ from the beginning and staying in the *i*th DI,

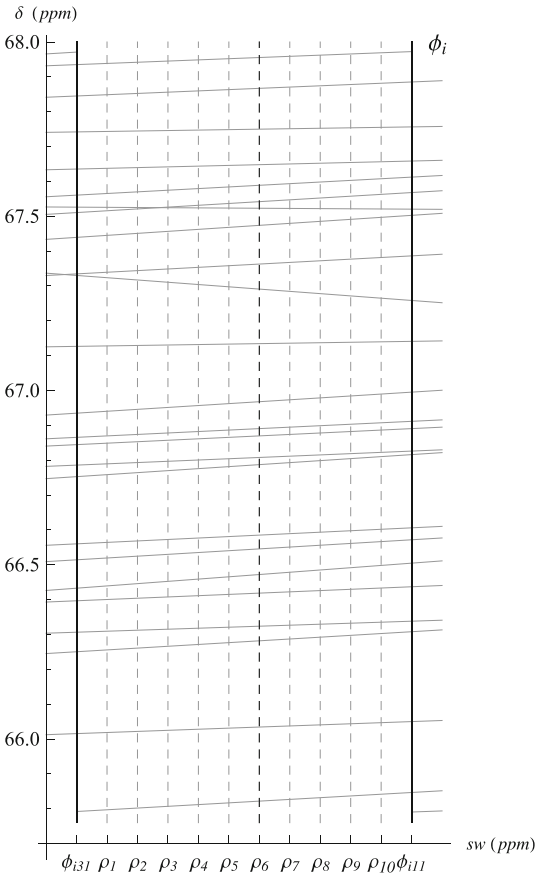

Fig. 3 Model of the ith Domain Interval ($\boldsymbol{\phi_i}$). The domain of the FF is divided into p intervals. Information for the ith DI is assigned to the respective element (ϕ_i) calculated previously (see Fig. 2 and Eq. 27), including the right endpoint (ϕ_{i11}), left endpoint (ϕ_{i31}), set of slopes (ϕ_{i2}) and the size of the interval (ϕ_{i32}). The DI is divided into d subintervals ($d = 11$ in the figure) giving $d - 1$ spectral options (see Eq. 28) represented by the vertical dashed lines labeled from ρ_1 to ρ_{10}. The aliasing processes of the CSs are represented by the oblique lines that correspond to the linear equations calculated using the CS values and the set of slopes stored in $\phi_{i2} : f_k = \delta_k + \phi_{i2k} * \rho$. The vertical dashed line labeled ρ_6, the best option calculated, corresponds to the spectrum shown in Fig. 6 (Sect. 4)

we obtain the n slope values by copying the set of slopes from the previous DI stored in list $\boldsymbol{M}$ with one difference; the value stored in ϕ_{i13} is assigned to the element μ_k, where k is equivalent to ϕ_{i12}. This part of the process is a loop [22] (Fig. 2f–h), where we assign the values of the n slopes to the second part of each element ($\boldsymbol{\phi}_{i2}$). In the same loop, the third part of each element (ϕ_{i3}) is calculated.

3.4 Exploring the DIs

Figure 3 shows the model of a DI, where ϕ_{i11} and ϕ_{i31} are the right endpoint and left endpoint of the ith DI, respectively. The DI is divided into d subintervals. The list $\boldsymbol{P}$ gives the set of the $d - 1$ resulting spectral options:

$$\boldsymbol{P} = \{\rho_h\}_{h=1}^{d-1}, \quad where \quad \rho_h = \phi_{i31} + h\frac{\phi_{i32}}{d}. \tag{28}$$

To obtain a criterion for assessing the convenience of a particular option, the observed chemical shifts (OCSs), list $\boldsymbol{\Theta}$ for each option listed in $\boldsymbol{P}$, are calculated from $\boldsymbol{\Delta}$ and from the slope list ϕ_{i2}:

$$\boldsymbol{\Theta} = \{\{\delta_k + \phi_{i2k} * \rho_h, \delta_k\}\}_{k=1}^{n}, \tag{29}$$

where n is the number of OCS values listed in decreasing numerical order. The list of the distances between the OCS values and their previous element on the list $\boldsymbol{\Theta}$ is defined by list $\boldsymbol{\Lambda}$:

$$\boldsymbol{\Lambda} = \{\lambda_k\}_{k=1}^{n}, \quad where\ \lambda_k = \begin{cases} \theta_{k1} - \theta_{(k+1)1} & 1 \le k \le n-1 \\ \rho_h - (\theta_{11} - \theta_{n1}) & k = n \end{cases} \tag{30}$$

where the last element has been introduced to the list in order to consider the OP's between θ_{11} and θ_{n1}.

Three terms judge the convenience of the current option (see "Appendix B"):

$$a)\ \Delta\theta min = Min[\boldsymbol{\Lambda}], \quad b)\ SImn = \frac{2\rho_h}{\Delta\theta min}\ and\ \ c)\ SImx = \frac{\rho_h}{DRmx}, \tag{31}$$

where $\Delta\theta min$ is the minimum shift difference between the OCSs in the list $\boldsymbol{\Lambda}$, and $DRmx$ is the greatest digital resolution allowed according to the maximum evolution period desired. At the end of this part, each element of $\boldsymbol{\Phi}$, in other words, each DI explored, gives a new list:

$$\boldsymbol{\Xi} = \{\{\rho_h, \boldsymbol{\Theta}, \{\Delta\theta min, SImn, SImx\}\}\}_{h=1}^{d-1} \tag{32}$$

The best option for a DI will be the ξ_h that has the greatest $\Delta\theta min$, since this option will be less demanding on DR. The second and third elements of the third part are calculated and used in the final steps to select a spectral width between different DIs, since these terms give the smallest and greatest total number of data points (SI) allowed, respectively. Figure 4a demonstrates the procedure for constructing the element ξ_h, and Fig. 4b illustrates the procedure for selecting the best option for a DI based on list $\boldsymbol{\Xi}$. The best option selected is an element ω_i from the final option list.

3.5 Making the list of options

The complete optimization method (Fig. 5) includes all of the procedures ($mk\Phi$ and $mk\omega$) described in the previous sections. First, given the CSs list ($\boldsymbol{\Delta}$), the ^{13}C resonance frequency in MHz (f) and the maximum evolution period in seconds ($t1$), we define the constants used in the method $n, cf, DRmx, lmt\ and\ d$. Second, $mk\Phi$ calculates the list of DIs(list $\boldsymbol{\Phi}$ in Fig. 5b). Third, for the ith DI, $mk\omega$ gives the data from the ith sw option (ω_i) and the list of the options is assigned to the list $\boldsymbol{\Omega}$ (Fig. 5c). The ith element of $\boldsymbol{\Omega}$ is the best option in the ith DI and the information stored in its three elements is used for establishing selection criteria in the following stages. Then, we select the options according to their $\Delta\theta min$ (Figs. 5d). The options selected must agree with the limit of resolution: $\Delta\theta min \ge 2DRmx$. Next, we calculate the elements ω_{i32} and ω_{i33} (minimum and maximum SI requirements) for each spectral option. Finally, we list the options in increasing order of ω_{i32}. The first element of the list $\boldsymbol{\Omega}$ is the best sw option.

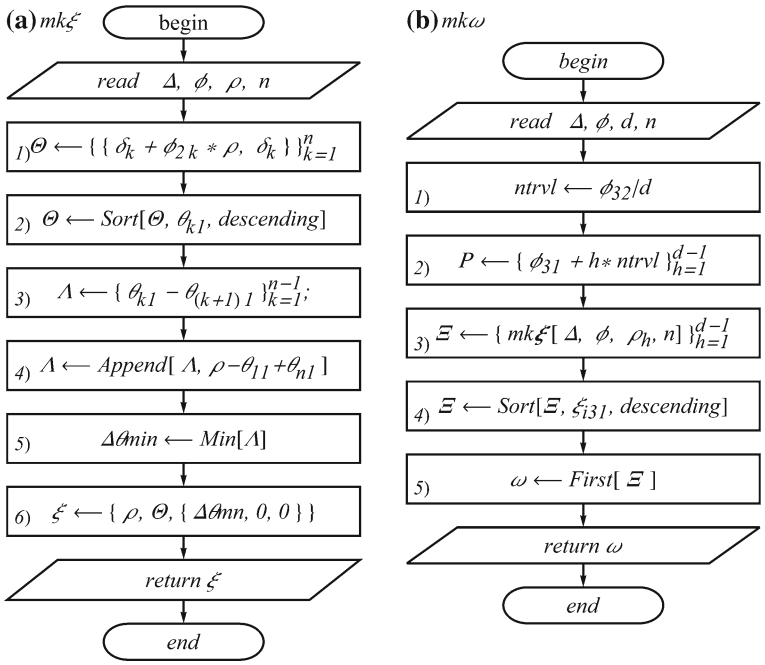

Fig. 4 Flowchart for the procedure of exploring the *i*th DI (**mkξandmkω**). Initially in **b**.*1*, the *i*th DI is divided into d subintervals. We then calculate the resulting set of $d-1$ spectral options (**b**.*2*). Each ρ_h is subsequently used to produce a list ξ_h by means of the procedure $mk\xi$ (**b**.*3* and **a**). For assessing ρ_h viability, $mk\xi$ calculates its list of OCSs (list $\boldsymbol{\Theta}$ in **a**.*1* and **a**.*2*) and its differences (list $\boldsymbol{\Lambda}$ in **a**.*3*). In order to consider all the OP's, $mk\xi$ appends a new element to the list $\boldsymbol{\Lambda}$ (**a**.*4*) and uses the new list to calculate $\Delta\theta min$ (**a**.*5*). The data from ρ_h is stored as a list of three elements and assigned to ξ_h (**a**.*6*). We assign the set of the lists ξ_h to $\boldsymbol{\Xi}$ (**b**.*3*), which is ordered in descending numerical order by comparing their $\Delta\theta min$ (**b**.*4*). Finally, we select the first element of the list $\boldsymbol{\Xi}$ (**b**.*5*) as the best option in the *i*th DI. This method is described in more detail in the Sect. 3.4

4 Results and discussion

Cholesterol is a good example for showing the practical use of this method. We apply the procedure to the list of CSs of the carbons with attached protons. This generates a list of 177 options, most of which are impractical for the high SI requirement. We chose the first two ranked elements, the options with the smallest SI requirements, 231 and 236, with spectral widths of 2.18074 and 2.40688 ppm, respectively (see Fig. 6) The Fourier transform algorithm is most suited to a number of data points that is raised to a power of two [16], and hence the SI calculated was rounded to 256.

Table 1 gives the calculated and experimental CSs for the first option. The number of significant figures is adjusted to reflect the resolution of the related spectra. The overlap of the closely situated ^{13}C resonances of C-7 and C-8 reported previously [12,1,18,15] is now resolved in the aliased HSQC spectra. The barely resolved signals for these carbons are separated by approximately 0.02 ppm. All the signals in the aliased HSQC spectrum of cholesterol acquired with the calculated parameters are unambiguously assigned.

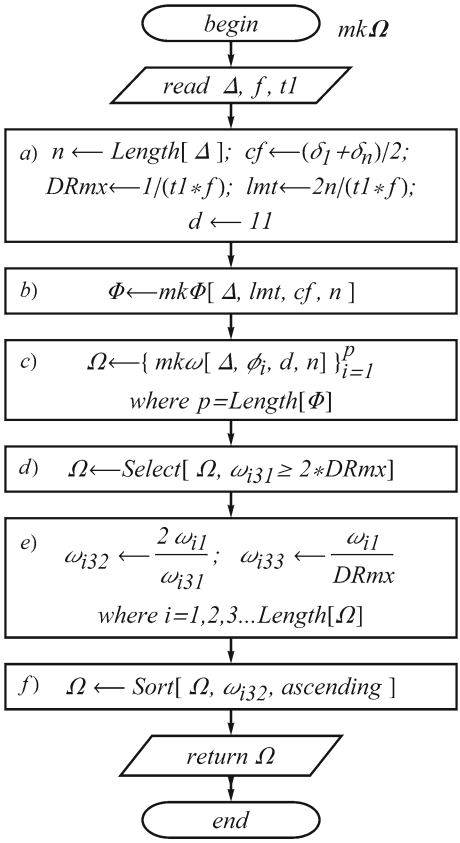

Fig. 5 Flowchart showing how to make the list of options (**mk*Ω***) First, from the CSs list (***Δ***), the ^{13}C resonance frequency in MHz (f) and the maximum evolution period in seconds ($t1$), we define the constants used in the procedures. We then calculate the DIs (the list ***Φ*** in **b**) and the list of *sw* options (the list ***Ω*** in **c**). The final three stages require selecting between the options from different DIs. The options selected must agree with the limit of resolution (**d**). We finally get the list of options ordered by their SI requirements (**f**). The first element of the list ***Ω*** is the best *sw* option. This method is described in more detail in the Sect. 3.5 and includes all of the procedures described in the previous sections

5 Conclusions

We can reduce the modulo function to a shorter expression. The definition of the PDs of the modulo function makes two things possible. First, we can divide the domain of the modulo function into DIs where the slopes that rule the aliasing of the n CSs considered are constants, and second, we can easily calculate these n slopes in each DI before calculating the OCSs. There is, therefore, no need to call the MF to search for good spectral options. This method is a viable alternative for optimizing the spectral widths in 2D NMR experiments.

6 Experimental

6.1 Compound and spectra

The cholesterol and solvents were purchased from Sigma-Aldrich. All experiments were carried out at 298 K using a 0.141-M solution in $CDCl_3$. The spectra of cholesterol were determined on a JEOL ECA 500 (^{1}H, 500.15992 MHz; ^{13}C, 125.7653 MHz)

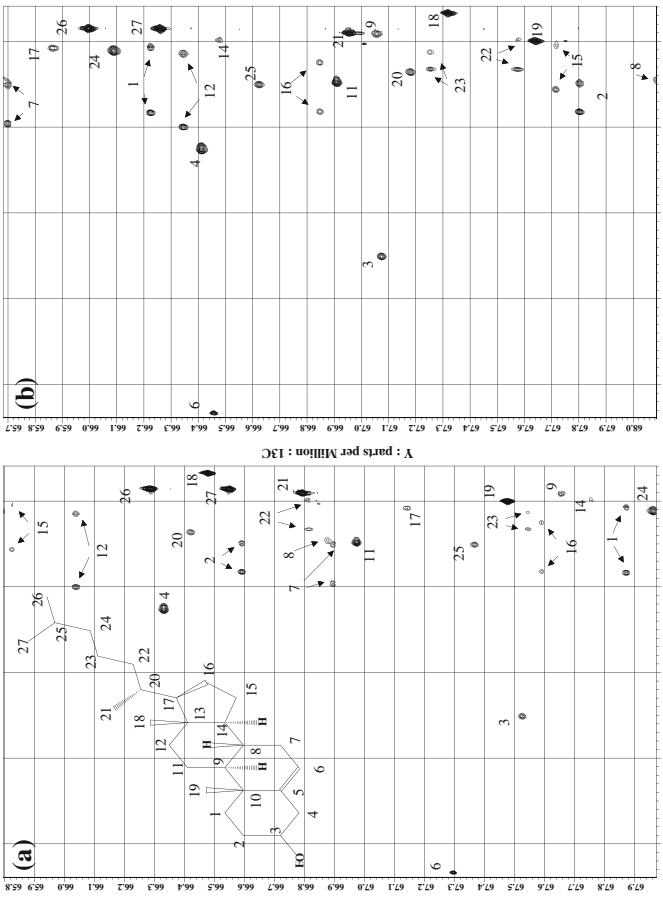

Fig. 6 Aliased HSQC spectra of cholesterol. Aliased HSQC spectrum of cholesterol, acquired with (**a**) ^{13}C sweep of 2.18074 ppm using 8 scans for each of the 256 increments and (**b**) ^{13}C sweep of 2.40688 ppm using 8 scans for each of the 256 increments. The ^{13}C aliased dimension was not linearly predicted but zero filled to 2K points. These two spectra correspond to the first two options calculated. The assignment of the spectra had the standard numbering

Table 1 Calculated and observed HSQC CSs using a spectral width of 2.18074 ppm and 256 data points

Position	δ_{Ha}	δ_{Hb}	δ_C	δ_C(Cal.)	δ_C(Exp.)	$\Delta\delta_C$(Cal. − Exp.)
C-1	1.07	1.84	37.34	67.87	67.87	0
C-2	1.49	1.83	31.69	66.59	66.59	0
C-3	3.51	–	71.88	67.52	67.52	0
C-4	2.25	–	42.34	66.32	66.33	−0.01
C-5	–	–	140.83	–	–	–
C-6	5.34	–	121.81	67.29	67.29	0
C-7	1.50	1.96	32.00	66.89	66.89	0
C-8	1.46	–	31.98	66.87	66.88	−0.01
C-9	0.92	–	50.20	67.65	67.66	−0.01
C-10	–	–	36.58	–	–	–
C-11	1.48	–	21.17	66.97	66.97	0
C-12	1.15	2.00	39.87	66.03	66.04	−0.01
C-13	–	–	42.40	–	–	–
C-14	0.98	–	56.85	67.75	67.75	0
C-15	1.04	1.56	24.39	65.82	65.83	−0.01
C-16	1.25	1.82	28.34	67.59	67.59	0
C-17	1.08	–	56.23	67.13	67.14	−0.01
C-18	0.67	–	11.95	66.47	66.48	−0.01
C-19	1.00	–	19.50	67.47	67.48	−0.01
C-20	1.36	–	35.89	66.42	66.42	0
C-21	0.90	–	18.81	66.79	66.79	0
C-22	0.99	1.33	36.28	66.81	66.81	0
C-23	1.13	1.33	23.93	67.54	67.55	−0.01
C-24	1.11	–	39.61	67.96	67.96	0
C-25	1.51	–	28.11	67.36	67.37	−0.01
C-26	0.86	–	22.67	66.28	66.29	−0.01
C-27	0.86	–	22.93	66.55	66.55	0

The number of significant figures for the results reflects the resolution of the related spectrum

spectrometer equipped with a 5-mm direct-detection probe with z gradients and controlled with Delta NMR Processing and Control Software [4].

The ^{13}C spectrum was acquired with an offset of 76.10877 ppm, sweep of 137.186 ppm, 32,768 data points, a digital resolution of 0.52653 Hz and 128 scans using the standard JEOL pulse sequence single-pulse decoupling. The HSQC parameters in the ^{1}H dimension had an offset of 2.96134 ppm, sweep of 6.78209 ppm, 2048 data points, and a digital resolution of 1.65631 Hz. The HSQC parameters in the ^{13}C dimension had an offset of 66.8815 ppm, sweep of 2.18074 ppm, and eight scans for each of the 256 increments were obtained using the JEOL pulse sequence based on pfg-HSQC CLUB (composite gradient leaving an undisturbed B_0 field) approach proposed by Hu and Shaka [10]. The ^{13}C aliased dimension was not linearly predicted but zero filled to 2K points. The spectra were processed without window functions in

F1. The unified scale [9] was used as a primary reference, with the ^{1}H resonance of TMS in dilute solution (volume fraction, $\varphi = 1\,\%$) in chloroform,$(CH_3)_4Si(\delta^1H = 0$, and $\delta^{13}C = 0$ for $\Xi^{13}C = 25.145\,020\,MHz$).

6.2 Computation

The CSs of carbons with attached protons were obtained from the ^{13}C spectrum and were selected using the Dept-135 spectrum. We put the ^{13}C CSs into the functions using four digits to the right of the decimal point, in the same way that they were provided by the acquisition program. The PDs $< lmt$ were not considered, and the DIs ≤ 0.001 ppm were not explored. A general line width is defined by cDRrq, where c=1. The calculations were performed with a set of five user-defined functions implemented in Wolfram Mathematica® [21] For Students (version 8.0.0.0) working with a Microsoft Windows (32-bit) platform (see Supplementary Material). The majority of figures were drawn with Mathematica® UDFs. except for Fig. 6. These two HSQC spectra were saved in eps format with Delta NMR processing and control software [4]

Acknowledgments J. F. Gómez-Reyes thanks Excelencia Profesional SENACYT-IFARHU Program, Panamá, for an awarded doctoral fellowship. M. A. Garcia-Ariza kindly made linguistic corrections, for which the authors express their gratitude.

7 Appendix A: List of acronyms

BIMF:	Build-in Mathematica® Function
CS:	Chemical Shift.
FF:	The Floor Function, Eq. (5).
HMBC:	Heteronuclear Multiple Bond Correlation.
HSQC:	Heteronuclear Single Quantum Coherence.
MF:	The Modulo Function (MF), Eq. (1).
OCS:	Observed Chemical Shift (chemical shift that is observed in an aliased spectrum).
PDs:	Points of Discontinuity.
TPPI:	Time Proportional Phase Increment (TPPI).
UDF:	User Defined Function.

8 Appendix B: List of variables

$cf(ppm)$:	Observation center. In this work it is defined as $cf = (\delta_1 + \delta_n)/2$.
d :	Number of subintervals into which the DIs was divided. In this work $d = 11$.
$\Delta\theta min(ppm)$:	Given a OCSs list (a list $\boldsymbol{\Theta}$), $\Delta\theta min$ is the smallest absolute value of the differences between the pairs of consecutive OCSs in the sorted list. Considering the possibility of doing a spectrum where $DRrq =$

	$DRmx$, we have $DRrq = DRmx = \Delta\theta min/2$. Therefore, to resolve all of the signals in such a spectrum, the value of $\Delta\theta min$ must be: $\Delta\theta min \geq 2 * DRmx$.
$DR\,(ppm)$:	Digital resolution. It is defined as: $DR = sw/SI$ [3].
$DRmx(ppm)$:	The greatest digital resolution that is allowed as stated by $t1 : DRmx = (1/t_1)/f$.
DRrq(ppm) :	The digital resolution that is required to resolve all of the signals in a spectrum. It is given in terms of the $\Delta\theta min$ and the factor c.Using $GLWs$ to describe the signals we gets: $\Delta\theta min = (c+1) * DRrq$. If we use $c = 1$, then $DRrq = \Delta\theta min/2$.
$f(MHz)$:	^{13}C resonance frequency.
GLW (ppm) :	A general line width used to define the line width of all of the signals in a spectrum. It is defined in terms of the $DRrq$ and the factor c: $GLW = c * DRrq$, where $c > 0$. So, the resolved resonances in a spectrum are separated by $\Delta\delta \geq (c+1)\,DRrq$.
h, i, j, k, l :	Subscript that are used to represent the elements of the lists.
$lmt(ppm)$:	Given a list of CSs (a list Δ), lmt is the smallest hypothetical sw that could be used to get a resolved spectrum. If we use $GLWs$ to describe the n signals, then each pair of consecutive resolved resonances in such a spectrum would be separated at least by $\Delta\delta = (c+1) * DRrq$. So, the size of lmt is given by $n * (c+1) * DRmx$. If we express the $DRmx$ in terms of $t1$ and use $c = 1$, then: $lmt = 2n/(t_1 * f)$.
n :	Number of CSs in the list $\boldsymbol{\Delta}$.
p :	Number of PDs, DIs, or sw options in the lists A, $\boldsymbol{\Phi}$ and $\boldsymbol{\Omega}$, respectively.
$SI(dimensionless)$:	The total number of data points.
$SImn\,(dimensionless)$:	Given a sw, $SImn$ is the smallest SI that is allowed as stated by $DRrq$. If we express the $DRrq$ in terms of $\Delta\theta\,min$, then: $SImm = sw/DRrq = 2 * sw/\Delta\theta min$.
$SImx\,(dimensionless)$:	Given a sw, $SImx$ is the greatest SI that is allowed as stated by $DRmx : SImm = sw/DRmx$.
$sw(ppm)$:	The observation range. Given a list of CSs (a list $\boldsymbol{\Delta}$), in a resolved aliased spectrum sw must be in the interval $lmt \leq sw \leq \delta_1 - \delta_n$
$t1(s)$:	The greatest evolution period that is allowed.

9 Appendix C: Representing the lists and their elements

In this paper, we use the structures known as lists [21,20] to collect the data together. The concept of list is analogous to the more familiar concepts of array [19] and matrix [2]. To deal with this kind of data object, we have used a notation similar to matrix notation: boldface capital letters to denote the lists and small letters, with some subscript indices to represent their elements surrounded by curly braces:

$$\boldsymbol{\Xi} = \left\{\xi_1, \xi_2, \ldots \xi_q\right\} \tag{33}$$

We can express the list $\boldsymbol{\Xi}$ in terms of a general element:

$$\boldsymbol{\Xi} = \{\xi_i\}_{i=1}^{q}$$

where ξ_i is used to denote the ith element of the list and the superscript q gives the length of the list. We note that we use exclusively Greek letters to denote a list or its elements and that the elements can be another list or a simple value. So ξ_i represents the ith element of the list $\boldsymbol{\Xi}$, ξ_{ij} represents the jth element of ξ_i, ξ_{ijk} represents the kth element of ξ_{ij}, and so on. For example, ξ_i can be defined as a list of three elements by the following way:

$$\boldsymbol{\Xi} = \left\{\{\xi_{11}, \xi_{12}, \xi_{13}\}, \{\xi_{21}, \xi_{22}, \xi_{23}\}, \ldots \{\xi_{q1}, \xi_{q2}, \xi_{q3}\}\right\} = \left\{\left\{\xi_{ij}\right\}_{j=1}^{3}\right\}_{i=1}^{q}$$

In the same way, the jth element of ξ_i can be a simple value, a list of lists or a simple list:

$$\boldsymbol{\Xi} = \left\{\left\{\xi_{i1}, \left\{\{\xi_{i2kl}\}_{l=1}^{2}\right\}_{k=1}^{n}, \{\xi_{i3k}\}_{k=1}^{3}\right\}\right\}_{i=1}^{q}$$

It is possible to express the general element with a formula, a procedure or an expression that includes the subscript indices:

$$\boldsymbol{A} = \left\{\left\{\{(4 + i * 2) * k/j - 4\}_{k=1}^{2}\right\}_{j=1}^{2}\right\}_{i=1}^{3}$$

The following is a summary of the lists that are used in this work. They are expressed in terms of general elements. The relations between the lists have been remarked on.

$\boldsymbol{\Delta} = \{\delta_i\}_{i=1}^{n}$: CSs list. The CSs are sorted in descending order. δ_i is used to calculate τ_i.

$\boldsymbol{T} = \left\{\left\{\tau_{ij}\right\}_{j=1}^{3}\right\}_{i=1}^{n}$: CSs distances data list. τ_i stores the data from the distance between δ_i to *cf.* τ_i is used to calculate the set of PDs of δ_i.

$\boldsymbol{E} = \left\{ \left\{ \{\varepsilon_{ijk}\}_{k=1}^{3} \right\}_{j=1}^{\frac{\tau_{i1}}{lmt}+\frac{1}{2}} \right\}_{i=1}^{Length[\boldsymbol{T}]}$: PDs data list. The data from the PDs are organized in sets according to their respective CSs. ε_i stores the data from the set of PDs of δ_i. ε_{ij} stores the data from the jth PD of δ_i.

$\boldsymbol{A} = \left\{ \{\alpha_{ij}\}_{j=1}^{3} \right\}_{i=1}^{p}$: PDs data list. The data from the PDs are organized into only one set. The elements of $\boldsymbol{A}$ correspond to all the lists ε_{ij} of $\boldsymbol{E}$ sorted in descending order by comparing their PD sizes. So, α_i stores the data from the ith PD. α_i is used to define the ith element in $\boldsymbol{\Phi}$.

$\boldsymbol{M} = \{\mu_k\}_{k=1}^{n}$: Slopes list. μ_k is the slope that rules the aliasing process of δ_k in the ith DI.

$\boldsymbol{\Phi} = \left\{ \left\{ \{\phi_{i1k}\}_{k=1}^{3}, \{\phi_{i2k}\}_{k=1}^{n}, \{\phi_{i3k}\}_{k=1}^{2} \right\} \right\}_{i=1}^{p}$: DIs data list. ϕ_i stores the data that is used for exploring the ith DI.ϕ_{i1} corresponds to α_i and ϕ_{i2} corresponds to $\boldsymbol{M}$. ϕ_{i3} stores the left endpoint value of the DI and the DI size. ϕ_i is used to define the ith element in $\boldsymbol{\Omega}$.

$\boldsymbol{P} = \{\rho_h\}_{h=1}^{q}$: DI *sw* options list. ρ_h is one of the q *sw*s options considered in a DI.

$\boldsymbol{\Theta} = \left\{ \{\theta_{kl}\}_{l=1}^{2} \right\}_{k=1}^{n}$: OCSs data list. Given a ρ_h, $\boldsymbol{\Theta}$ is the list of its OCSs. θ_{k1} stores the kth OCS and θ_{k2} stores the kth CS. The elements are sorted in descending order by comparing the θ_{k1} part.

$\boldsymbol{\Lambda} = \{\lambda_k\}_{k=1}^{n}$: OCSs differences list. $\boldsymbol{\Lambda}$ lists the $n-1$ differences between each of the two consecutive OCSs listed in $\boldsymbol{\Theta}$. λ_n is included in order to consider the OP's between θ_{11} and θ_{n1}.

$\boldsymbol{\Xi} = \left\{ \left\{ \xi_{h1}, \left\{ \{\xi_{h2kl}\}_{l=1}^{2} \right\}_{k=1}^{n}, \{\xi_{h3k}\}_{k=1}^{3} \right\} \right\}_{h=1}^{q}$: DI *sw* options data list. Given the ith DI and its corresponding list $\boldsymbol{P}$, ξ_h stores the data used for assessing the ρ_h viability. ξ_{h1} corresponds to ρ_h and ξ_{h2} corresponds to $\boldsymbol{\Theta}$. ξ_{h3} stores the data used as a criterion to select the best option from the q elements

$$\boldsymbol{\Omega} = \left\{ \left\{ \omega_{i1}, \left\{ \{\omega_{i2kl}\}_{l=1}^{2} \right\}_{k=1}^{n}, \left\{ \omega_{i3k} \right\}_{k=1}^{3} \right\} \right\}_{i=1}^{p} :$$

of $\boldsymbol{P}$. The best option in the ith DI corresponds to the ith element in $\boldsymbol{\Omega}$. *sw* options data list. ω_i corresponds to the best *sw* option in the ith list $\boldsymbol{\Xi}$. ω_{i31} stores the $\Delta\theta min$ calculated from ω_{i2}. $\boldsymbol{\Omega}$ is finally sorted in descending order by comparing the ω_{i31} part. So, the first element of this rearranged list is the best option calculated by the method.

Table 2 UDFs used in the flowcharts

	UDF	Auxiliary UDF	lists made by the function	Output lists
(a)	$mk\Phi[\Delta, lmt, cf, n]$	none	$\boldsymbol{T}, \boldsymbol{E}, \boldsymbol{\Phi}, \boldsymbol{M}$	$\boldsymbol{\Phi}$
(b)	$mk\xi[\Delta, \phi_i, \rho_i, n]$	none	$\boldsymbol{\Theta}, \boldsymbol{\Lambda}, \xi_i$	ξ_i
(c)	$mk\omega[\Delta, \phi_i, d, n]$	mkξ	$\boldsymbol{P}, \boldsymbol{\Xi}, \omega_i$	ω_i
(d)	$mk\Omega[\Delta, f, t1]$ (main function)	$mk\Phi, mk\omega$	$\boldsymbol{\Phi}, \boldsymbol{\Omega}$	$\boldsymbol{\Omega}$

The UDF arguments are enclosed by square brackets

Table 3 BIMFs, signs and operators used in the flowchart expressions

	Examples	Resultant lists
a)	$\boldsymbol{A} \leftarrow \left\{ \left\{ \{(4 + i * 2) * k/j - 4\}_{k=1}^{2} \right\}_{j=1}^{2} \right\}_{i=1}^{3}$	$\boldsymbol{A} = \{\{\{2, 8\}, \{-1, 2\}\}, \{\{4, 12\}, \{0, 4\}\}, \{\{6, 16\}, \{1, 6\}\}\}$
b)	$\mathbf{B} \leftarrow$ Flatten[**A**, first level]	$\boldsymbol{B} = \{\{2, 8\}, \{-1, 2\}, \{4, 12\}, \{0, 4\}, \{6, 16\}, \{1, 6\}\}$
c)	$\mathbf{K} \leftarrow$ Sort[**B**, β_{i2}, descending]	$\boldsymbol{K} = \{\{6, 16\}, \{4, 12\}, \{2, 8\}, \{1, 6\}, \{0, 4\}, \{-1, 2\}\}$
d)	$\mathbf{O} \leftarrow$ Select[**K**, $\kappa_{i2} \leq 6$]	$\boldsymbol{O} = \{\{1, 6\}, \{0, 4\}, \{-1, 2\}\}$
e)	$\mathbf{Z} \leftarrow \{\text{First}[o_i]\}_{i=1}^{\text{Length}[\mathbf{O}]}$	$\boldsymbol{Z} = \{1, 0, -1\}$
f)	$\mathbf{X} \leftarrow \{\text{Sign}[\zeta_i]\}_{i=1}^{\text{Length}[\boldsymbol{Z}]}$	$\boldsymbol{X} = \{1, 0, -1\}$
g)	$\mathbf{N} \leftarrow \{\text{Abs}[\chi_i]\}_{i=1}^{\text{Length}[\boldsymbol{X}]}$	$\mathbf{N} = \{1, 0, 1\}$
h)	$\mathbf{H} \leftarrow$ Append[**N**, 15]	$\mathbf{H} = \{1, 0, 1, 15\}$
i)	$\mathbf{I} \leftarrow$ {Min[**H**]}	$\mathbf{I} = \{0\}$

The BIMFs arguments are given in a simplified manner and enclosed by square brackets. a) The equality sign (=) means equality in the usual sense of the world. A left arrow (←) ( ) is used to indicate that a value or a list of values is stored in a variable. The arithmetic operators for addition (+), subtraction (-), multiplication (*) and division (/) are the same used in C, Basic and Fortran expressions. b) Flatten[*list, first level*] deletes the inner braces of *list* at the *first level*. c) Sort[*list, part, descending*] arranges the elements of *list* in *descending* order by comparing a specific *part* of each element. d) Select[*list, criterion*] picks out all elements of *list* for which *criterion* is true. e) First[*list*] and Length[*list*] give the first element and the number of elements in *list,* respectively. f) Sign[x] gives -1, 0 or 1 depending on whether the real number x is negative, zero, or positive. g) Abs[x] gives the absolute value of the real number x. h) Append[*list, element*] gives *list* with *element* added to its end. i) Min[*list*] gives the minimum of *list*

10 Appendix D: Flowchart expressions

The method of optimization is divided into four procedures. Each of these procedures is explained in detail with the aid of a flowchart and has a corresponding UDF (Online Resource 1). The UDF names $mk\Phi$, $mk\xi$, $mk\omega$, $mk\Omega$ (see Table 2) and the BIMF names [21] *Flatten, Sort, Select, First*, *Length, Sign, Abs,* and *Min* (see Table 3) has been used with the aim of introducing the four UDFs that compute each stage of the method.

References

1. J.W. Blunt, J.B. Stothers, Org. Magn. Reson. **9**, 439 (1977)
2. W. Brown, *Matrices and Vector Spaces* (Marcel Dekker, Inc., New York, 1991)
3. T. Claridge, *High-Resolution NMR Techniques in Organic Chemistry* (Pergamon, Oxford, 1999)
4. Delta NMR Processing and Control Software, Version 4.3.6. (Jeol USA, Inc., Peabody, 2006)
5. A.J. Dunn, P.J. Sidebottom, Magn. Reson. Chem. **43**, 124 (2005)
6. J.F. Gómez-Reyes, A. Ariza-Castolo, Analysis of the folding functions and a dynamic approach to the problem of optimizing spectral widths in 2D experiments. *53rd Experimental Nuclear Magnetic Resonance Conference* (2012). http://www.enc-conference.org/portals/0/flipbook/. Accessed 26 June 2012
7. J.F. Gómez-Reyes, A. Ariza-Castolo, Analysis of the folding functions and a dynamic approach to the problem of optimizing spectral widths in the 13C dimension of heteronuclear 2D experiments. *53rd Experimental Nuclear Magnetic Resonance Conference* (2012). http://www.enc-conference.org/portals/0/Abstracts2012/ENC20121291.1250VER.1.pdf. Accessed 26 June 2012
8. R.L. Graham, D.E. Knuth, O. Patashnik, *Concrete Mathematics* (Addison-Wesley, Reading, 1988), pp 67–69, 81–83
9. R.K. Harris, E.D. Becker, R. Goodfellow, P. Granger, Magn. Reson. Chem. **40**, 489 (2002)
10. H. Hu, A.J. Shaka, J. Magn. Reson. **136**, 54 (1999)
11. D. Jeannerat, in *Encyclopedia of Magnetic Resonance*, ed. by R.K. Harris, R.E. Wasylishen, (Wiley, Chichester, 2011), Published online. doi:10.1002/9780470034590.emrstm1187
12. D. Jeannerat, Magn. Reson. Chem. **38**, 415 (2000)
13. D. Jeannerat, Magn. Reson. Chem. **41**, 3 (2003)
14. D. Jeannerat, J. Magn. Reson. **186**, 112 (2007)
15. P. Joseph-Nathan, G. Mejía, D. Abramo-Bruno, J. Am. Chem. Soc. **101**, 1289 (1979)
16. J. Keeler, *Understanding NMR Spectroscopy* (Wiley, Chichester, 2005)
17. L. Leithold, *The Calculus With Analytic Geometry*, 4th edn. (Harper & Row, New York, 1981)
18. H.H. Mantsch, I.C.P. Smith, Can. J. Chem. **51**, 1384 (1973)
19. L. Nyhoff, *C++: An Introduction to Data Structures* (Prentice Hall, Upper Saddle River, 1999)
20. P. Wellin, S. Kamin, R. Gaylord, *An Introduction to Programming with Mathematica*, 75, 3rd ed. (Cambridge University Press, Cambridge, 2005), pp 53–59
21. Wolfram Mathematica® 8 For Students, version 8.0.0.0. (Wolfram Research, Inc., Champaign, 2010)
22. T. Worth, *Basic For Everyone* (Prentice Hall, Englewood Cliffs, 1976)