

## Defining the sampling space in multidimensional NMR experiments: What should the maximum sampling time be?

Thomas Vosegaard\*, Niels Chr. Nielsen

Center for Insoluble Protein Structures (inSPIN), Interdisciplinary Nanoscience Center (iNANO) and Department of Chemistry, University of Aarhus, Langelandsgade 140, DK-8000 Aarhus C, Denmark

### ARTICLE INFO

#### Article history:

Received 22 December 2008

Revised 6 April 2009

Available online 17 April 2009

#### Keywords:

Sampling

Multidimensional NMR spectroscopy

Solid-state NMR

Macroscopically aligned samples

Complex error function

Apodization

Truncation

### ABSTRACT

Efficient sampling of signals is a key issue for multiple-dimensional NMR experiments to establish the best ratio between experiment time and spectral quality. Focussing on the most widely used sampling strategy using standard rectangular sampling and data analysis by Fourier transformation, a central question is concerned with determining the optimal maximum sampling time in the individual dimensions. The spectral resolution depends directly on this choice, as do the overall experiment times when addressing the indirect dimensions. We present a theoretical, numerical, and experimental analysis of the sampling space problem and propose approaches to efficient sampling for typical cases.

© 2009 Elsevier Inc. All rights reserved.

### 1. Introduction

Long experiment times are one of the great obstacles for multi-dimensional NMR experiments, typically induced by intrinsic low sensitivity of the sample, the need for transient-consuming phase cycling [1,2], or because it is necessary to sample a large number of points in indirect dimensions to achieve sufficient spectral resolution. To reduce the experiment time, the costly achievement of high resolution in indirect dimensions is often compromised by choosing a shorter maximum sampling time,  $t_s$ , corresponding to fewer time increments. In addition, this dilemma has inspired a host of different strategies for sampling the indirect dimensions of multiple-dimensional experiments to optimize efficiency and resolution through appropriate selection of data points between times 0 and  $t_s$ .

To illustrate the maximum sampling-time problem, we consider the conventional and by far most used sampling technique with data points arranged in a rectangular grid in the range  $0 \leq t \leq t_s$ . This sampling scheme is convenient since it allows for easy data processing using standard Fourier transformation procedures. Generally, it will be safe to treat each spectral dimension independently, although certain schemes, e.g., circular [3] or triangular [3,4] sampling, may introduce correlation between the involved spectral dimensions for severely truncated datasets.

In this paper, we present a systematic analysis of the spectral consequences of the chosen size of the sampling space, here defined by the maximum sampling time,  $t_s$ . To keep the analysis as transparent as possible and to avoid correlation of spectral dimensions, we restrict the present study to equidistant sampling, which – although an increasing number of exotic new sampling methods is introduced and extensively debated – still represents by far the most used sampling strategy in the NMR community. This by itself is a strong motivation for our present focus. In this context, we note that although alternative sampling strategies relying on equidistant sampling may induce minor correlation between sampling in different dimensions, as discussed above, the results may still be used as guidelines for estimating the overall geometry of the sampling space as expressed by the choice of  $t_s$ . Likewise, we should note that the present analysis cannot be used to predict the resolution and sensitivity when using more sophisticated processing models such as maximum entropy [5–7], where the resolution depends on the deconvolution kernel and the abundance of peak artefacts is more of an issue than the sensitivity/SNR depending on the white noise in the FID. The manuscript is organized as follows. First, we introduce the basic concepts of sampling, resolution, sensitivity, and through this the theory underlying our analysis of sampling strategies. Next, we introduce recipes for optimal sampling under various conditions. Finally, we demonstrate the power of efficient sampling as compared to conventional sampling in the context of  $^1\text{H}$ – $^{15}\text{N}$  separated-local-field (SLF) experiments for a  $^{15}\text{N}$  labeled single crystal of *N*-acetyl-L-leucine (NAL) serving as a model system for uniaxially oriented membrane proteins.

\* Corresponding author. Fax: +45 8619 6199.  
E-mail address: [tv@chem.au.dk](mailto:tv@chem.au.dk) (T. Vosegaard).

## 2. Theory

Our analysis of efficient sampling strategies is conducted in several steps, which – along with some of the necessary definitions concerning parameters for the NMR signal, our target spectral quality, and our manipulations, all of which are summarized in Table 1 – may be listed as: (i) assuming a certain *natural linewidth* of the resonances (induced by relaxation, field inhomogeneity, pulse sequence imperfections, etc.),  $\lambda_n$ , we calculate the signal-to-noise ratio (SNR) and sensitivity (*vide infra*), the resulting linewidth (henceforth denoted the *target linewidth*,  $\lambda_t$ ), and the severity of the truncation-induced sinc wiggles (identified by the *wiggle height*,  $w$ ) as a function of  $t_s$ . (ii) We consider apodization as a means to reduce wiggles and increase the sensitivity by introducing the *apodization linewidth*,  $\lambda_a$ , to the analysis. (iii) While  $t_s$  and  $\lambda_a$  in combination determine the spectral appearance, it is more useful to describe the spectral quality based on the sensitivity, the target linewidth, and the wiggle height and use these sample-dependent parameters to determine suitable values for  $t_s$  and  $\lambda_a$ . That is, the abundance and proximity of resonances may define a need for a certain spectral resolution (upper limit for  $\lambda_t$ ), and similarly, maximum tolerable wiggle height is sample dependent since it depends on, e.g., the risk of falsely interpreting the wiggles as real peaks. In order to use these findings as guidelines to set up of the sampling space in multidimensional NMR experiments, we present empirical formulas to calculate  $t_s$  and  $\lambda_a$  as functions of  $\lambda_t$  and  $w$ . (iv) Finally, we demonstrate the relevance of efficient sampling in the context of oriented-sample solid-state NMR spectroscopy, although here emphasizing that our analysis and sampling strategies is completely general and apply to all kinds of multiple-dimensional experiments.

### 2.1. Sampling with optimal signal-to-noise ratio

The challenge of choosing the right size of the sampling space for a particular sample is illustrated by the one-dimensional example in Fig. 1. Fig. 1a shows the SNR as a function of the sampling time,  $t_s$  for Lorentzian and Gaussian peaks with a natural (full-width-half-height; FWHH) linewidth  $\lambda_n$ , under the assumption of equidistant sampling. From this graph, it is clear that a sampling time of  $t_s \approx 0.4/\lambda_n$  (for a Lorentzian resonance;

$t_s \approx 0.52/\lambda_n$  for a Gaussian resonance) would yield maximum SNR. It is also evident, however, from the corresponding FID and spectrum shown in Figs. 1e and f that this is achieved at the expense of substantial broadening and appearance of truncation wiggles. A quantification of these undesired effects is given in Figs. 1c and d, which report the linewidth and wiggle height (i.e., maximum amplitude of sinc wiggles relative to the intensity of the resonance) in the spectrum as a function of  $t_s$ . From these graphs we may conclude, that the truncation effects approximately double the linewidth of a Lorentzian line when sampled for maximum SNR.

While optimizing the SNR is relevant for acquisition of the direct dimension in cases where the repetition delay is long compared to  $t_s$  (i.e., the overall experiment time does not depend on our choice of  $t_s$ ), we will normally seek to optimize the SNR per unit time when the total experiment time is proportional to  $t_s$ . This is the case for all indirect dimensions using equidistant sampling. Ernst et al. [8] define the SNR per unit time, typically referred to as the *sensitivity*

$$\text{sensitivity} = \frac{\text{SNR}}{\sqrt{t_s}}. \quad (1)$$

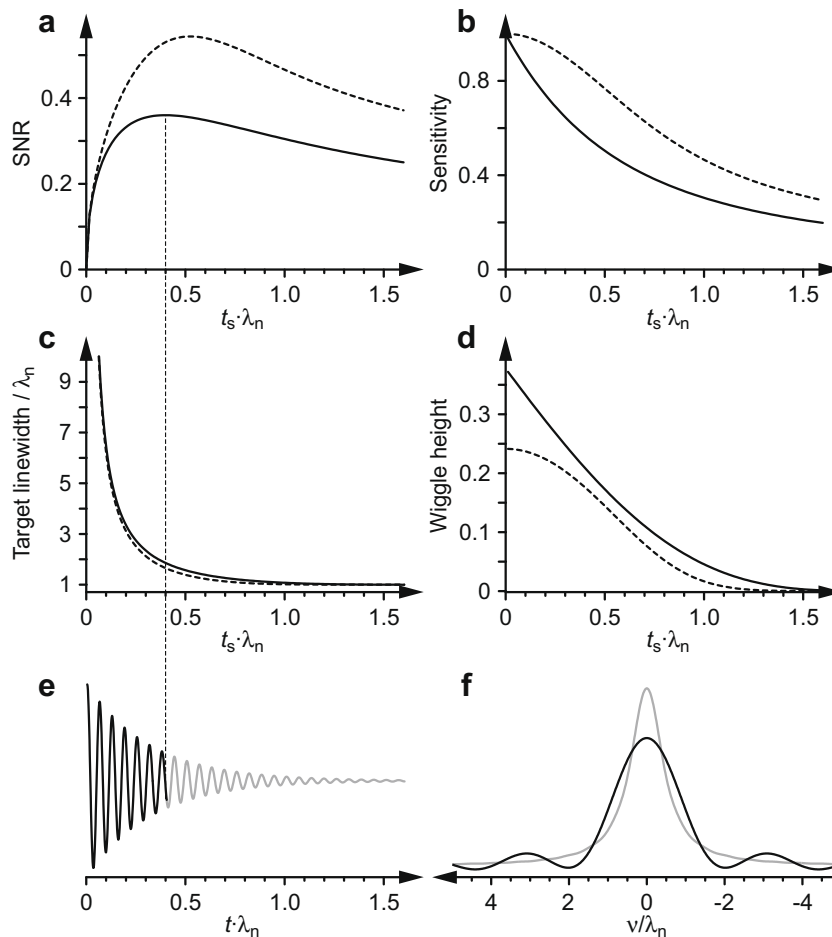
The plot of the *sensitivity* (Fig. 1b) reveals a monotonously decreasing functionality, always favouring the shortest possible  $t_s$ .

Considering these observations together, it is in general impossible to maximize the SNR or *sensitivity* without compromising the spectral quality. The excessive truncation linebroadening may be undesirable for resolution purposes, and wiggles may be as intense as  $\sim 35\%$  of the peak height (see Fig. 1d), implying that they may in unfavourable cases be incorrectly interpreted as signals or cause undesired cancellation of real signals. The severity of these artefacts is sample related, since the problems caused by truncation linebroadening and wiggles depend on, e.g., the desired resolution, the density of resonances, and the intrinsic SNR. For example, if the sample is characterized by only few, well-separated resonances, the resolution may not be an issue. Likewise, if we have a very low SNR, we need not worry too much about wiggles, which will anyway be hidden in the noise or reduced by apodization. On the other hand, if we are seeking weak cross-peaks close to an intense diagonal peak, wiggles from the diagonal may hamper a correct spectral interpretation.

**Table 1**

List of the various parameters used in the article.

Parameter	Description
$t_s$	Maximum sampling time. Denotes the maximum sampling time for a particular dimension of a multidimensional NMR experiment, either generically or detailed by a specific dimension ( $n$ )
$\lambda$	Linewidth. Generic representation of a full-width at half-height (FWHH) linewidth
$\lambda_n$	Natural linewidth. Denotes the FWHH linewidth for resonances in a particular dimension of a multidimensional NMR experiment. This linewidth is induced by relaxation, field inhomogeneity, pulse sequence imperfections, etc
$\lambda_a$	Apodization linewidth. Denotes the FWHH linewidth of the apodization function (Lorentzian or Gaussian)
$\lambda_t$	Target linewidth. Denotes the resulting FWHH linewidth for a resonance with a natural linewidth $\lambda_n$ , under specific acquisition and processing conditions
$w$	Wiggle height. Denotes the fractional height (relative to the peak height) of the most intense sinc wiggle resulting from truncation of the time-domain signal.
$S$	Signal. Denotes the signal intensity (in arbitrary units) of either the time- or frequency-domain signal
$N$	Noise. The noise is assumed to be white noise entering the time-domain signal as a random function within a certain amplitude
SNR	Signal-to-noise ratio. The SNR is defined on an arbitrary scale where the signal intensity is defined as $S(v=0)$ in Eq. (5) and the noise is defined in Eq. (13)
nSNR	Normalized SNR. Denotes the SNR achieved with a specific set of parameters divided by the optimal SNR corresponding to $t_s \rightarrow \infty$ and $\lambda_a = \lambda_n$
<i>sensitivity</i>	Sensitivity. The <i>sensitivity</i> represents the SNR per unit time (cf. Eq. (1)) and is used to measure the efficiency of the sampling of indirect dimensions, where the overall acquisition time is proportional to $t_s$
$A_L$	Lorentzian apodization function. Defined in Eq. (3)
$A_G$	Gaussian apodization function. Defined in Eq. (3)
$L$	Lorentzian linewidth. $L = \lambda_a + \lambda_n$
$G$	Gaussian linewidth. $G = \sqrt{(\lambda_a^2 + \lambda_n^2)/(4 \ln 2)}$
$G_a$	Shorthand notation for $G_a = \lambda_a / \sqrt{2 \ln 2}$
$\Delta$	Doublet splitting parameter. Measure of the “valley” between two close peaks
$np$	Number of points. Number of complex data points in FID/spectrum
$sw$	Spectral width
$\Delta v$	Digital resolution. Denotes the distance between two datapoints in the frequency domain of a discrete spectrum, i.e., $sw/np$



**Fig. 1.** Plots of (a) signal-to-noise ratio (SNR), (b) sensitivity, (c) linewidth (FWHH), and (d) wiggle height for Lorentzian (solid lines) and Gaussian (dashed lines) resonances as a function of the maximum sampling time,  $t_s$ , the lineshape, and the natural linewidth,  $\lambda_n$ . (e) Free-induction decay for a Lorentzian resonance, and (f) corresponding frequency-domain signal (black line,  $t_s = 0.4/\lambda_n$ ; grey line,  $t_s = 1.6/\lambda_n$ ).

The above example raises two questions that need to be addressed before proceeding with the detailed analysis: (i) can we use the target linewidth as a measure for the spectral resolution? and (ii) can we use a continuous description, which allows us to gain more analytical insight than does a discrete description?

## 2.2. Measure of spectral resolution

To test the resolution obtained using specific sampling parameters, we evaluate the peak splitting parameter [6]:  $\Delta = 1 - 2C/(A + B)$ , where A and B are the intensities of two peaks and C is the intensity in the valley between them. A peak splitting parameter of 0 means no resolution, while two completely separated parameters have a peak splitting parameter of 1. Figs. 2a and b show the domains where  $\Delta > 0.1$  in shade in plots of the peak distance vs. the sampling time. For comparison, these graphs also include plots of the target linewidths, which show the same trends as the peak splitting parameters. In particular, plots of  $1.1 \cdot \lambda_t$  (Lorentzian resonances) and  $1.15 \cdot \lambda_t$  (Gaussian resonances) are virtually identical to the condition  $\Delta = 0.1$ , showing that for truncated data, the target linewidth is indeed an ideal measure for the spectral resolution.

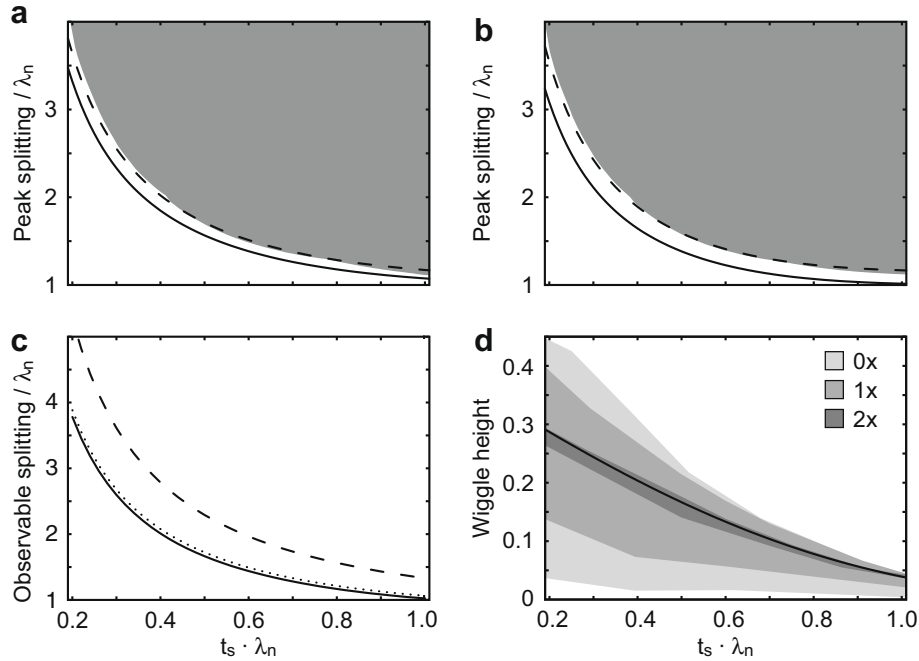
## 2.3. Discrete and continuous representations

The next question concerns how well a continuous spectral description represents the discrete representation. The characteris-

tics of a continuous Fourier transformation correspond to a discrete description with infinite zero-filling. Thus, it is the amount of zero-filling that distinguishes the discrete description from the continuous. To investigate the spectral consequences of zero-filling, Fig. 2c shows the resolution power (taken as the condition  $\Delta = 0.1$  for Lorentzian peaks) with and without zero-filling. Here, we note the expected trend that interpolation by zero-filling substantially improves the resolution. Another aspect of the interpolation is the amount of truncation wiggles, here expressed as the maximum difference between the spectrum resulting from processing of the truncated dataset and a spectrum with an ideal Lorentzian peak with the target linewidth. The wiggle intensity varies significantly depending on the peak position relative to the data bins in the spectrum. Fig. 2d illustrates these effects (for Lorentzian peaks), and the shaded areas represent typical domains in which the wiggle height falls for various degrees of zero-filling. We observe that the typical approach of ample zero-filling provides the most predictable behaviour, and that all cases show similar trends, so the discrete case (corresponding to infinite zero-filling) will provide a good measure for the wiggle height.

## 2.4. Lorentzian and Gaussian resonances

In the following, we consider a resonance for which the free-induction decay (FID) is decaying exponentially leading to either a Lorentzian or Gaussian lineshape with a natural linewidth  $\lambda_n$ . This linewidth needs not be the relaxation-induced linewidth,



**Fig. 2.** (a and b) Grey areas represent domains where the doublet peak splitting ( $\Delta$ ) is larger than 0.1 for Lorentzian (a) and Gaussian (b) resonances in plots of the peak splitting as a function of  $t_s$  employing zero-filling to  $2t_s$ . For comparison, the target linewidth,  $\lambda_t$  (solid lines) and  $X\lambda_t$  (dashed lines),  $X = 1.1$  and  $X = 1.15$  for Lorentzian and Gaussian resonances, respectively. (c) Plot of the minimum observable peak splitting ( $\Delta = 0.1$ ) as a function of  $t_s$  employing no zero-filling (dashed line) and 1 time zero-filling (doubling of the number of points) (solid line) for a Lorentzian resonance. (d) Wiggle height as a function of the zero-filling factor. 0 means no zero-filling, 1 means doubling of the number of points, i.e., zero-filling to  $np \cdot 2^n$  points where  $n = 0.1, \text{ and } 2$ . The shaded areas represent the typical interval of the wiggles, but the actual amount depends on the resonance frequency of the peaks relative to the position of the data points. The solid line represents the wiggle height in the case zero-filling to  $np \cdot 2^3$  or more points.

but may contain contributions from field inhomogeneities, and pulse-sequence imperfections. It reflects the linewidth achieved when the signal is sampled to convergence, i.e., the spectrum contains no truncation artefacts. For simplicity we will assume that the signals are on-resonance in the rotating frame giving the following expressions for the FIDs

$$\begin{aligned} S_L'(t; \lambda_n) &= \exp(-\pi\lambda_n t) \\ S_G'(t; \lambda_n) &= \exp\left(-\frac{(\pi\lambda_n t)^2}{4 \ln 2}\right) \end{aligned} \quad (2)$$

In analogy to this, we may introduce Lorentzian and Gaussian apodization functions, with apodization linewidth,  $\lambda_a$ , as

$$\begin{aligned} A_L(t; \lambda_a) &= \exp(-\pi\lambda_a t) \\ A_G(t; \lambda_a) &= \exp\left(-\frac{(\pi\lambda_a t)^2}{4 \ln 2}\right) \end{aligned} \quad (3)$$

The apodized signal results from multiplying the apodization function to the ideal FID, leading to a convolution of the Fourier transformed functions in the frequency domain. We will restrict the analysis to Lorentzian apodization of Lorentzian signals and Gaussian apodization of Gaussian signals, since the concept of *matched filtering* predicts that the optimum SNR is achieved by apodizing the signal with an apodization function of the same functionality [8]. Likewise, we will not consider any resolution-enhancing filtering functions, since this is beyond the scope of this paper, where we aim at finding conditions for maximizing the SNR. Clearly, all resolution-enhancing functions will increase the contribution from regions of the time-domain signal with low SNR implying a decrease in the overall SNR.

The apodized FIDs are given by

$$\begin{aligned} S_L(t, \lambda_a; \lambda_n) &= \exp(-\pi(\lambda_a + \lambda_n)t) = \exp(-\pi Lt) \\ S_G(t, \lambda_a; \lambda_n) &= \exp\left(-\frac{(\lambda_a^2 + \lambda_n^2)(\pi t)^2}{4 \ln 2}\right) = \exp(-(\pi G t)^2) \end{aligned} \quad (4)$$

where we introduced the shorthand notations  $L = \lambda_a + \lambda_n$  and  $G = \sqrt{(\lambda_a^2 + \lambda_n^2)/(4 \ln 2)}$ . If the signals are sampled in the time interval  $0 \leq t \leq t_s$ , the real part of the frequency-domain signals are given by the Fourier transformations of the apodized FIDs:

$$\begin{aligned} S_L(v, t_s, \lambda_a; \lambda_n) &= \text{Re}\left\{\int_0^{t_s} S_L(t; \lambda_n) \exp(-i2\pi vt) dt\right\} \\ &= \frac{L + (2v \sin(2\pi vt_s) - L \cos(2\pi vt_s))e^{-\pi Lt_s}}{\pi(4v^2 + L^2)} \\ S_G(v, t_s, \lambda_a; \lambda_n) &= \text{Re}\left\{\int_0^{t_s} S_G(t; \lambda_n) \exp(-i2\pi vt) dt\right\} \\ &= \frac{\exp(-(v/G)^2)}{2\sqrt{\pi}G} \text{Re}\{\text{Erf}(\pi G t_s + iv/G)\}, \end{aligned} \quad (5)$$

where  $\text{Erf}(z)$  represents the complex error function.

The frequency-domain plots in Fig. 1f are described by the Lorentzian formula in Eq. (5). We note that in the limit  $t_s \rightarrow \infty$ , the wiggles disappear and the expressions approach the expected Lorentzian and Gaussian frequency-domain functions:

$$\begin{aligned} \lim_{t_s \rightarrow \infty} (S_L(v, t_s, \lambda_a; \lambda_n)) &= \frac{L}{\pi(4v^2 + L^2)} \\ \lim_{t_s \rightarrow \infty} (S_G(v, t_s, \lambda_a; \lambda_n)) &= \frac{\exp(-(v/G)^2)}{2\sqrt{\pi}G} \end{aligned} \quad (6)$$

The intensity of the signal in the frequency domain is given by the signal intensity at  $v = 0$ :

$$S_L(v=0, t_s, \lambda_a; \lambda_n) = \frac{1 - \exp(-\pi L t_s)}{\pi L} \quad (7)$$

$$S_G(v=0, t_s, \lambda_a; \lambda_n) = \frac{\text{Erf}(-\pi G t_s)}{2\pi G},$$

which for both Lorentzian and Gaussian resonances are monotonously increasing functions.

### 2.5. Target linewidth

To include sample issues into our discussion, a critical parameter is the linewidth we can accept for our experiment, i.e., the so-called *target linewidth*  $\lambda_t$ , which in our context is larger than or equal to the natural linewidth  $\lambda_n$ . We define the target linewidth as the actual linewidth of the truncated and apodized frequency-domain signal in Eq. (5). To calculate it, we use the expressions normalized signal

$$S^n(v, t_s, \lambda_a; \lambda_n) = \frac{S(v, t_s, \lambda_a; \lambda_n)}{S(v=0, t_s, \lambda_a; \lambda_n)} \quad (8)$$

and find the target linewidth, which satisfies the equation

$$S^n(v = \lambda_t/2, t_s, \lambda_a; \lambda_n) = \frac{1}{2}. \quad (9)$$

There is no analytical solution to this equation unless in the limit of infinitely long sampling, but since we have found that the wiggles will never be as high as 50%, there is only one solution for positive values of  $\lambda_t$ , which is easy to find numerically, thereby enabling us to determine the entity  $\lambda_t(\lambda_s, \lambda_a; \lambda_n)$ .

### 2.6. Wiggles

We define the normalized wigggle function as the difference between the expression of the normalized frequency domain signal in Eq. (7) and the ideal Lorentzian/Gaussian shapes with the same target linewidth:

$$W_L(v, t_s, \lambda_a; \lambda_n) = S_L^n(v, t_s, \lambda_a; \lambda_n) - \frac{1}{1 + (2v/\lambda_t(t_s, \lambda_a; \lambda_n))^2}$$

$$W_G(v, t_s, \lambda_a; \lambda_n) = S_G^n(v, t_s, \lambda_a; \lambda_n) - \exp(-4\ln 2 (v/\lambda_t(t_s, \lambda_a; \lambda_n))^2) \quad (10)$$

The maximum wigggle height,  $w$ , is then the maximum absolute value of the wigggle function

$$w = \text{Max}_v |W(v, t_s, \lambda_a; \lambda_n)| \quad (11)$$

This expression for the wigggle height, when transferred to a discrete sampling space, corresponds to the case of infinite zero-filling, since the Fourier transform in Eq. (5) goes to infinite time. We note that wiggles are virtually absent in a non-zero-filled spectrum, but this comes at the expense of severe lack of resolution since each resonance is defined by very few (typically one) data points. The distance between two data points in the frequency domain without zero-filling will be  $\Delta v = 1/t_s$ , and since truncation only appears in the region of short sampling times ( $t_s < 1/\lambda_n$ ), the approximate number of points per resonance may be approximated by  $\Delta v/\lambda_n < 1$ .

### 2.7. Signal-to-noise ratio and sensitivity

White noise enters in our description as a random function in each sampling point, and the overall noise will be proportional to the square root of the number of sampling points. While data sampling obviously is conducted in a discrete manner, we will remain with a continuous description. With the restriction of equidistant sampling, we may simply define the random noise as  $N(t_s) = \sqrt{t_s}$ . Here we should remember, that the noise appears in the “raw” sig-

nal before apodization, but when the signal is apodized, the noise reduces, and our estimate the total noise in the FID is the rms average of the apodized noise

$$N(t_s, \lambda_a) = \sqrt{\int_0^{t_s} (A(t, \lambda_a))^2 dt}. \quad (12)$$

For Lorentzian and Gaussian resonances, respectively, the explicit expressions become

$$N_L(t_s, \lambda_a) = \begin{cases} \sqrt{t_s} & \text{for } \lambda_a = 0 \\ \sqrt{\frac{1 - \exp(-2\pi\lambda_a t_s)}{2\pi\lambda_a}} & \text{for } \lambda_a > 0 \end{cases} \quad (13)$$

$$N_G(t_s, \lambda_a) = \begin{cases} \sqrt{t_s} & \text{for } \lambda_a = 0 \\ \sqrt{\frac{\text{Erf}(-2\pi G_a t_s)}{2\pi G_a}} & \text{for } \lambda_a > 0 \end{cases}$$

with the shorthand notation  $G_a = \lambda_a/\sqrt{2\ln 2}$ . Measures of the SNR and the *sensitivity* are given by

$$\text{SNR}(t_s, \lambda_a; \lambda_n) = \frac{S(v=0, t_s, \lambda_a; \lambda_n)}{N(t_s, \lambda_a)}, \quad (14)$$

$$\text{sensitivity}(t_s, \lambda_a; \lambda_n) = \frac{S(v=0, t_s, \lambda_a; \lambda_n)}{N(t_s, \lambda_a)\sqrt{t_s}},$$

where the expressions for the signal and noise are given in Eqs. (7) and (13), respectively. When comparing the SNR of different experiments, it may be useful to compare the normalized SNR, which reaches a maximum value of 1, and which may be calculated as

$$n\text{SNR}(t_s, \lambda_a; \lambda_n) = \frac{\text{SNR}(t_s, \lambda_a; \lambda_n)}{\text{SNR}(t_s \rightarrow \infty, \lambda_a = \lambda_n; \lambda_n)} = \begin{cases} \text{SNR}(t_s, \lambda_a; \lambda_n)\sqrt{\lambda_n}/2\pi & \text{Lorentzian} \\ \text{SNR}(t_s, \lambda_a; \lambda_n)\sqrt{\lambda_n}/\sqrt{\ln 2} & \text{Gaussian} \end{cases} \quad (15)$$

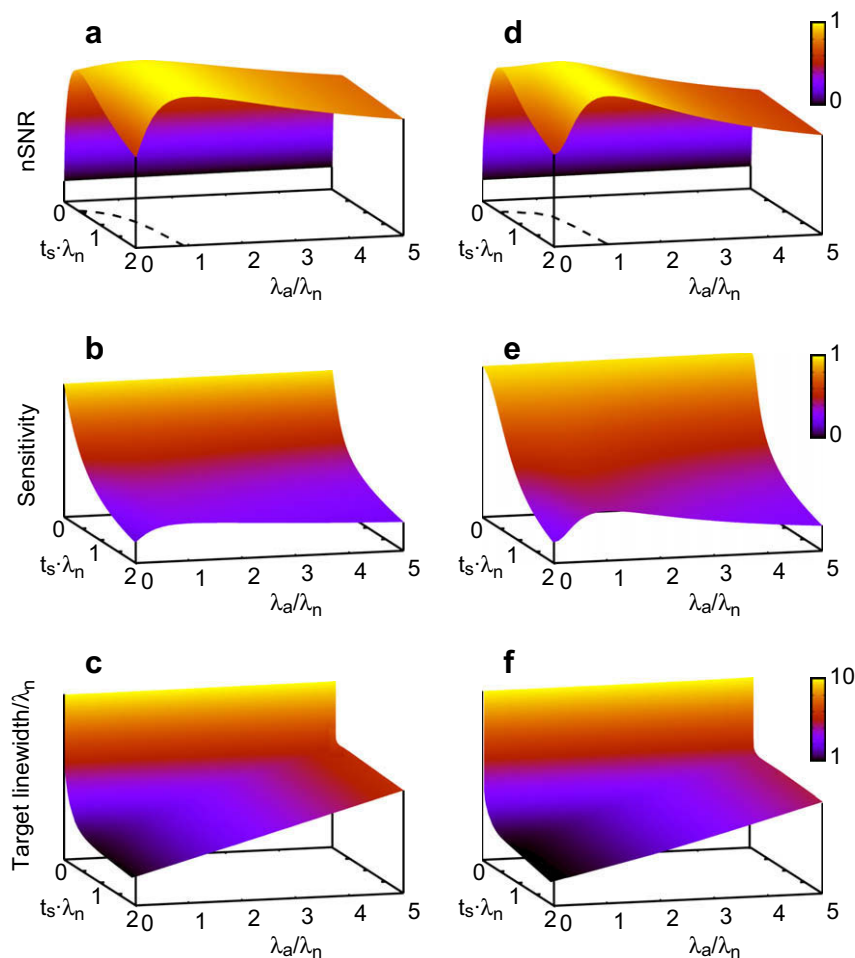
where  $\text{SNR}(t_s \rightarrow \infty, \lambda_a = \lambda_n; \lambda_n)$  represents the maximum achievable SNR, corresponding to an infinitely long sampling time with matched filtering.

## 3. Results and discussion

Depending on what is the most critical issue for a given experiment, it is the interplay between sampling and data treatment that is our handle to optimize the quality of a dataset obtained within a given time frame. The quality of a spectrum will be expressed in terms of SNR, sinc wiggles, and target linewidth. In the following, we will address these aspects in more detail leading to general recipes for (i) analyzing the spectral outcome for a particular choice of sampling parameters ( $t_s$  and  $\lambda_a$ ) and (ii) defining the optimum sampling space and the accompanying data processing for practical application in the indirect dimension(s) of multi-dimensional NMR experiments.

### 3.1. The apodization linewidth

The use of apodization may be desirable to increase the SNR or to reduce truncation effects, as demonstrated in Fig. 1. To extend these results and to exploit the 2D space spanned by the user-specified parameters,  $t_s$  and  $\lambda_a$ , we have calculated the SNR, the *sensitivity*, and the target linewidth as a function of  $t_s$  and  $\lambda_a$ , as illustrated in Fig. 3. Focussing first on the SNR (Figs. 3a and d), it is interesting to note that these functions, for apodization linewidths in the region  $\lambda_a < \lambda_n$ , have local maxima at short values of  $t_s$ , as illustrated by dashed lines below the surfaces, with the slices along  $t_s$  for  $\lambda_a = 0$  corresponding to the plots in Fig. 2a. The global maxima for these graphs are found at the condition  $\lambda_a = \lambda_n$ , while more than 90% of the maximum SNR is reached for  $t_s \approx 0.7/\lambda_n$  in the Lorentzian case and  $t_s \approx 0.6/\lambda_n$  in the Gaussian case. We note that the



**Fig. 3.** Plots of (a and d) the normalized SNR, (b and e) *sensitivity*, and target linewidth,  $\lambda_t$  (c and f) as a function of the maximum sampling time,  $t_s$ , and apodization linewidth,  $\lambda_a$ , for Lorentzian (a–c) and Gaussian (d–f) resonances. The dashed lines in (a) and (d) represent the local SNR maxima for the particular apodization linewidths.

$\lambda_a = \lambda_n$  condition providing the maximum SNR is known as matched filtering [8].

A plot of the *sensitivity* (Figs. 3b and e) is less interesting since it shows monotonically decreasing function as a function of  $t_s$  for all values of  $\lambda_a$ , favouring the shortest possible value for  $t_s$ . However, when inspecting the plots of the resulting linewidths, here named the *target linewidth* (Figs. 3c and f), we observe a very significant increase in the target linewidth for small values of  $t_s$ , as already discussed for the special case  $\lambda_a = 0$  in relation to Fig. 1. We observe that for all values of  $\lambda_a$  (at least in the considered region  $0 < \lambda_a < 5$ ) the truncation-contribution to the target linewidth is important in the regime of short  $t_s$ .

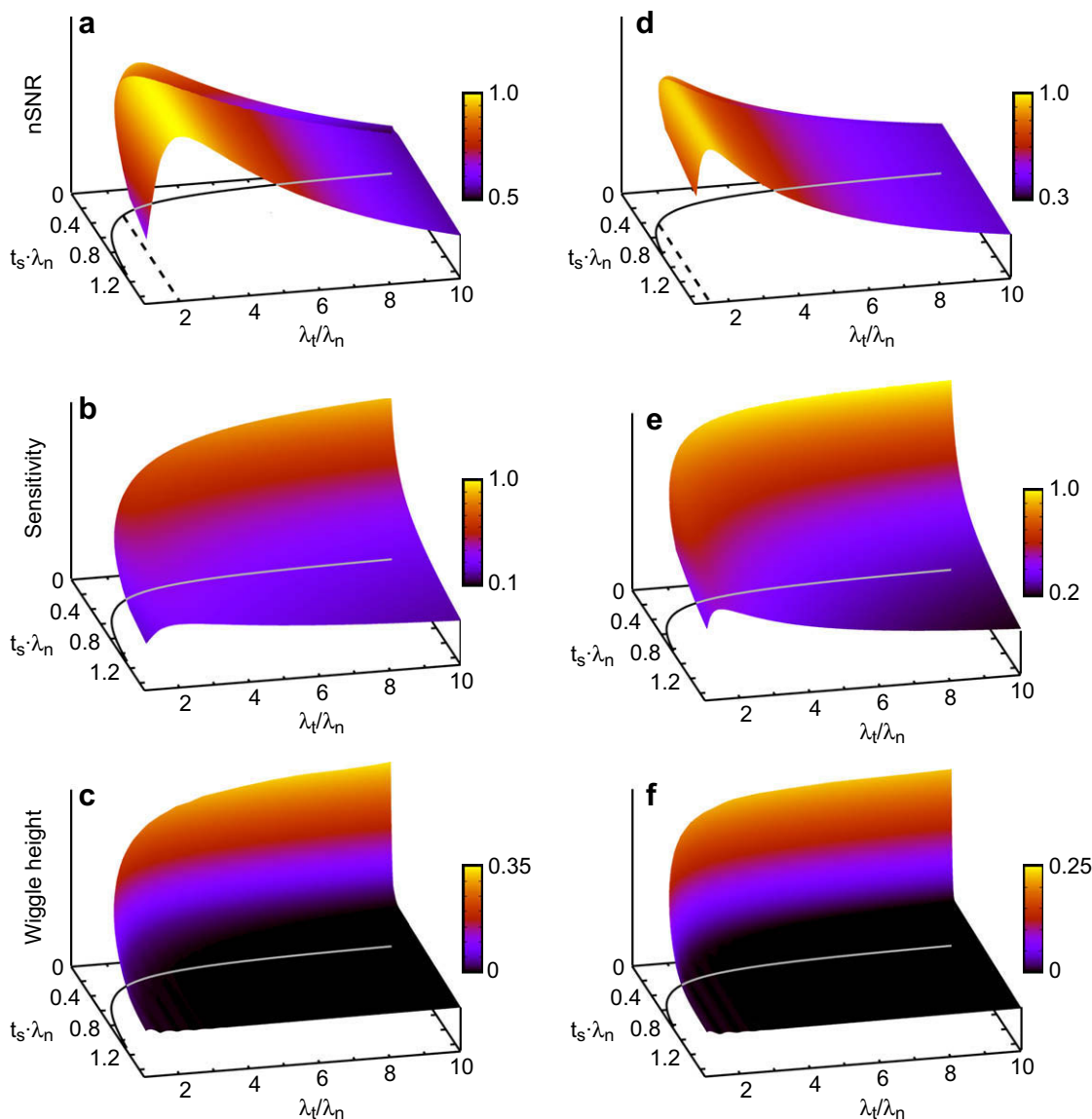
### 3.2. The target linewidth

For the spectroscopist, the apodization is an important tool used to increase the quality of the spectrum, but as evidenced by the plots in Figs. 3c and f, there is only a simple linear relation between  $\lambda_a$  and  $\lambda_t$  in the regime of long sampling, while this linearity does not exist in the truncation regime. Therefore, and because it is the *target linewidth* that defines the appearance of the spectrum, it seems relevant also to map the spectral properties such as SNR, *sensitivity*, and sinc wiggle height as a function of the target linewidth instead of the apodization linewidth.

Turning to a description of the SNR, *sensitivity*, and wiggle height as a function of  $t_s$  and  $\lambda_t$  (Fig. 4) reveals plots that are markedly different from those in Fig. 3. Probably the most striking dif-

ference is that the functions are not defined for small values of  $t_s$ . This is reflecting the fact that when  $t_s$  is decreased, at some point the truncation linewidth will exceed the target linewidth. In Fig. 4, the line in the zero-contour plane indicates this cut-off. We note that this curve corresponds to  $\lambda_a = 0$  and therefore is identical to the plot of the linewidth as a function of  $t_s$  reported in Fig. 1c.

Regarding the SNR and *sensitivity*, the conclusions made from the plots in Fig. 3 are not changed by the new representation shown in Fig. 4. For the SNR, we observe a global maximum at  $\lambda_t = 2\lambda_n$  for a Lorentzian resonance and  $\lambda_t = \sqrt{2}\lambda_n$  for a Gaussian resonance, corresponding to matched filtering in both cases. When the sample and experimental conditions permit sampling in the SNR regime, which is characterized by the fact that the overall acquisition time is not influenced by the choice of  $t_s$ , we will just choose a sufficiently long acquisition time to avoid any truncation artefacts. This will normally be the case for the direct dimensions in solid-state NMR, where the acquisition time ( $t_s$ ) is small compared to the repetition delay. The maximum *sensitivity* is still achieved at the shortest possible value for  $t_s$ , but now with the lower bound indicated by solid lines in Fig. 4. This lower bound is achieved when  $\lambda_a = 0$ , and hence it is the truncation linewidth that defines  $\lambda_t$ . Consequently, wiggles will make a substantial spectral signature close to this condition as evidenced by the plots of the wiggle height (Figs. 4c and f). In the worst case, the wiggle height is as large as 35% for a Lorentzian resonance (25% for a Gaussian resonance), potentially imposing severe problems in terms of



**Fig. 4.** Plots of the (a and d) normalized SNR, (b and e) sensitivity, and wiggle height,  $r$  (c and f) as a function of the maximum sampling time,  $t_s$ , and target linewidth,  $\lambda_t$ , for Lorentzian (a–c) and Gaussian (d–f) resonances. The solid lines below the surfaces represent the cutoff value for  $t_s$  below which the data is undefined. The dashed lines in (a) and (d) represent the SNR maxima.

interpretation as false peaks. To avoid this, sampling conditions using slightly longer values for  $t_s$  should be used, since the wiggle height drops down quite rapidly as  $t_s$  increases. In the following, we will devise a recipe for choosing the right sampling parameters.

### 3.3. Analyzing the consequence of specific sampling parameters

In the following, we provide a fast way to analyze the spectral outcome for a given choice of  $t_s$  and  $\lambda_a$ . First, we will revisit the results of the truncation-induced linebroadening (Fig. 1c) and wiggle height (Fig. 1d). While the quantities SNR and sensitivity are readily described analytically by Eq. (14),  $\lambda_t$  and  $w$  cannot be described by simple analytical formulas. We have used the software Mupad [9] to numerically determine the wiggle height and target linewidth as reported in Fig. 4, and subsequently found the following empirical formulas describing these entities

$$\lambda(t) \approx 1 + \frac{a}{\exp(bt) - 1} \quad (16)$$

$$w(t) \approx a \cdot \text{Erfc}(bt)^c \quad (17)$$

which, with the coefficients  $a$ ,  $b$ , and  $c$  given in Table 2, represent the curves with less than 5% deviation.  $\text{Erfc}(x)$  denotes the complementary error function  $\text{Erfc}(x) = 1 - \text{Erf}(x)$ . Note that, we have introduced a normalized linewidth ( $\lambda$ ) and sampling time ( $t$ ), which in the case of no apodization simply are the normalized expressions of  $\lambda_t$  and  $t_s$ :

$$\begin{aligned} \lambda &= \lambda_t / \lambda_n \quad \text{and} \\ t &= t_s \cdot \lambda_n \end{aligned} \quad (18)$$

If apodization is applied, the normalized linewidth is related to the sampling parameters by the following expression

$$\lambda = \begin{cases} \frac{\lambda_t}{\lambda_a + \lambda_n} & \text{(Lorentzian)} \\ \frac{\lambda_t}{\lambda_a^2 + \lambda_n^2} & \text{(Gaussian)} \end{cases} \quad (19)$$

while the expression in Eq. (16) still holds for  $t_s$ . Equipped with these expressions, it is very simple to analyze evaluate the target linewidth (Eqs. (18) and (15)), wiggle height (Eq. (16)), SNR, and sensitivity (Eq. (14)), from a specific choice of  $t_s$  and  $\lambda_a$ .

In order to facilitate the analysis, we have made a utility available from our web site (<http://www.bionmr.chem.au.dk>). This util-

**Table 2**

Coefficients for the empirical description of the truncation linewidth ( $\lambda$ ) and wiggle height ( $w$ ).<sup>a</sup>

Lineshape	Parameter	<i>a</i>	<i>b</i>	<i>c</i>	Equation
Lorentzian	$\lambda$	1.669	2.729		(16)
	$w$	0.374	1.141	1.063	(17)
	$w$	1.867	1.525		(21)
Gaussian	$\lambda$	2.291	3.735		(16)
	$w$	0.241	1.519	2.461	(17)
	$w$	2.466	1.947		(21)

<sup>a</sup> c.f. Eqs. (16), (17), and (21).

ity does not rely on the approximate expressions in Eqs. (15) and (16), but uses the more accurate values from the plots in Fig. 1.

**3.4. Optimized sampling conditions**

Having established the basis for analyzing specific acquisition and apodization parameters as described above, it is very relevant bring the analysis one step further and develop a simple recipe for setting up the optimized sampling conditions, in terms of the maximum sampling time, in order to reduce the experiment time or achieve the desired resolution. To facilitate this, the flow chart in Fig. 5 provides a guideline to determination of the relevant experimental parameters under certain conditions. The case handled in the previous section assumed that the natural linewidth was known, and analyzed the effect of certain acquisition- and processing parameters, thus ending up in box 1 in Fig. 5. In the following, we will demonstrate when the remaining boxes are relevant.

Considering altogether the sample- and experiment-specific requirements for resolution and truncation artefacts given a cer-

tain natural linewidth, we have identified the following two entries as the most efficient way to take into account all the sample-related issues:

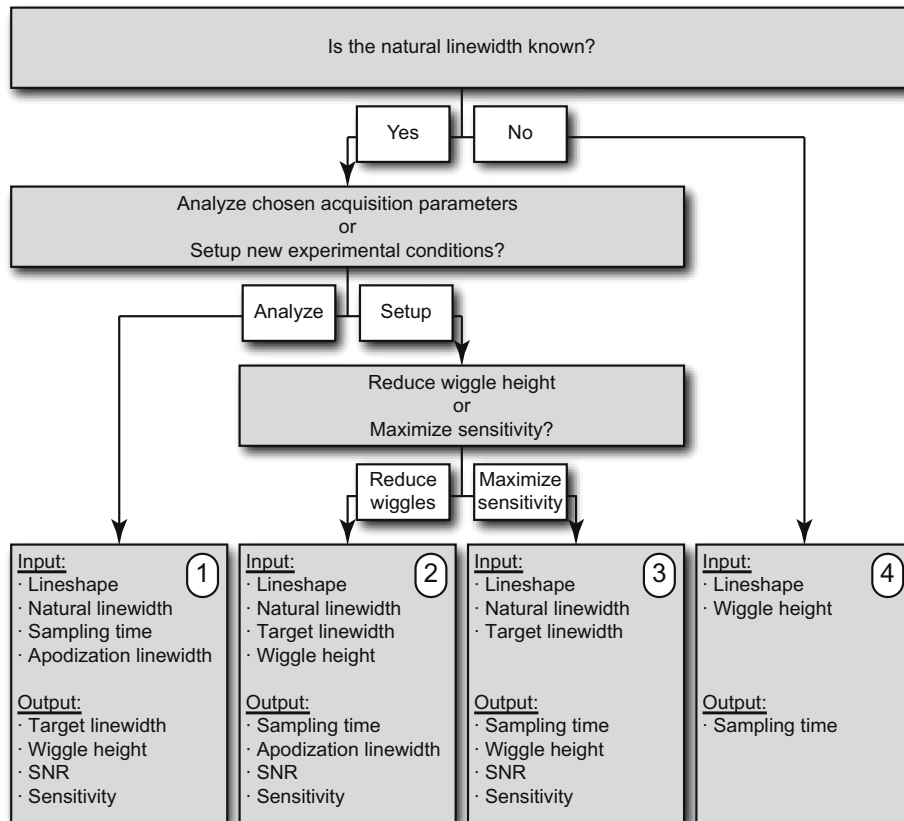
- Specify the maximum acceptable target linewidth ( $\lambda_t$ ) to provide necessary resolution.
- Specify the maximum acceptable wiggle height ( $w$ ) to avoid erroneous interpretations of sinc wiggles.

With these choices, we follow the flow chart and end up in box 2 in Fig. 5. Based on these two parameters ( $\lambda_t$  and  $w$ ), we may readily derive the optimized sampling conditions by a simple analysis of the graphs in Fig. 4. We want to extract values for  $t_s$ ,  $\lambda_a$ , the SNR, and the *sensitivity* for the specified values of  $\lambda_t$  and  $w$ . This procedure involves (i) finding the value for  $t$  from  $w$  using Eq. (15). While it is not possible to obtain an analytical solution, it is simple to find the solution numerically. (ii) Using this value for  $t$  as input to Eq. (15) allows us to calculate  $\lambda$ , which represents the truncation-contribution to the target linewidth. (iii) From this, we calculate the remaining linewidth,  $\lambda_r = \lambda_t / \lambda$ . (iv) If  $\lambda_r < \lambda_n$ , there exists no solution (*vide infra*), but when  $\lambda_r \geq \lambda_n$ ,  $\lambda_a$  is given by

$$\lambda_a = \begin{cases} \lambda_r - \lambda_n & \text{(Lorentzian)} \\ \sqrt{\lambda_r^2 + \lambda_n^2} & \text{(Gaussian)} \end{cases} \quad (20)$$

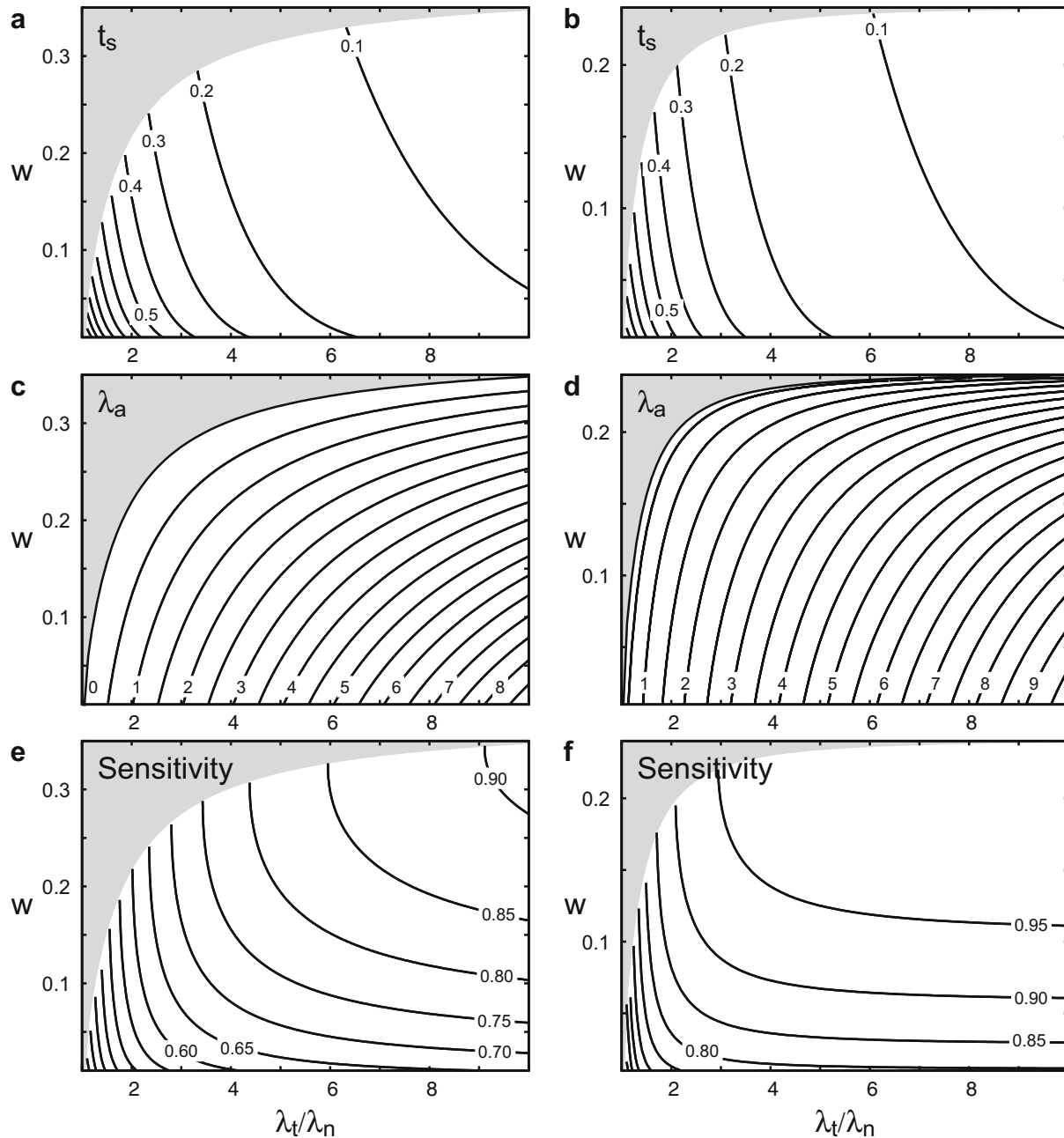
$$t_s = \frac{t}{\lambda_r}$$

Equipped with these parameters, the SNR and *sensitivity* may readily be calculated from Eq. (14). The failing condition ( $\lambda_r < \lambda_n$ ) occurs when the truncation is not too severe, in which case the wiggle height does not reach so high values (see the low  $\lambda_t$ -region in Figs. 4c and f).



**Fig. 5.** Flow chart illustrating the general procedure for analyzing or setting up experimental conditions. The four boxes marked 1, 2, 3, and 4 refer to different procedures to determine the unknown sampling and processing parameters.





**Fig. 6.** Contour plots showing the maximum sampling time (a and b), the apodization linewidth (c and d), and the sensitivity (e and f) as a function of the setup parameters  $\lambda_t$  and  $w$ , for Lorentzian (a, c, and e) and Gaussian (b, d, and f) resonances. The contour levels are equidistant in all plots. The grey region in the upper left corner represents an undefined region.

A simpler alternative to the approach above would be to go for the highest possible sensitivity, i.e., by sampling in the regime where the target linewidth is only given by truncation (corresponding to  $\lambda_a = 0$ ). Following the flow chart this leads us to box 3 in Fig. 5. In this case, the only free variable is the target linewidth, while  $t_s$  is found from Eq. (16). Knowing these two parameters, the remaining parameters are determined as described for the analysis above.

To ease the setup of proper experimental conditions, Fig. 6 shows contour plots of  $t_s$ ,  $\lambda_a$ , and the sensitivity as a function of  $\lambda_t$  and  $w$ , as a condensed representation of the data in Fig. 3 and Fig. 4. Equipped with these graphs it is straightforward to setup the experiments to fulfill the requirements on  $\lambda_t$  and  $w$  and rapidly evaluate the prize paid in terms of eventually lost sensitivity. In addition, the two procedures for setting up new experimental

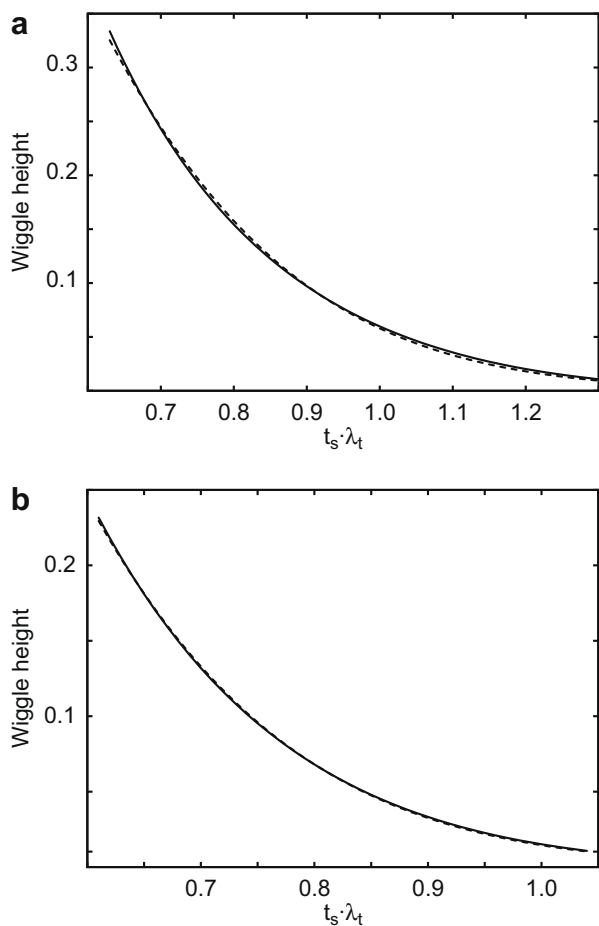
protocols are implemented in the above-mentioned utility available from our web site (<http://www.bionmr.chem.au.dk>).

As a final point to the analysis, we note that there is a simple relation between the product of the sampling time and the target linewidth and the wobble height as reported in Fig. 7. These curves may be described by the approximate empirical formulas

$$w(t_s, \lambda_t) \approx a \cdot \text{Erfc}(bt_s \lambda_t), \quad (21)$$

where the coefficients  $a$  and  $b$  are listed in Table 2.

The plots in Fig. 7 are particularly useful if the natural linewidth is unknown, as illustrated in box 4 in Fig. 5. In this case,  $t_s$  may be determined from the decided values for  $w$  and  $\lambda_t$  and then apodize the spectrum appropriately to match the target linewidth. The appropriate apodization linewidth may be calculated as



**Fig. 7.** Plot of the wiggle height as a function of the product  $t_s \cdot \lambda_t$  for (a) Lorentzian and (b) Gaussian lineshapes. Solid lines represent the values achieved from the data in Fig. 4, while the dashed lines represent the functions in Eq. (22).

$$\lambda_a = \begin{cases} \lambda_t - \lambda_e & \text{(Lorentzian)} \\ \sqrt{\lambda_t^2 + \lambda_e^2} & \text{(Gaussian)}, \end{cases} \quad (22)$$

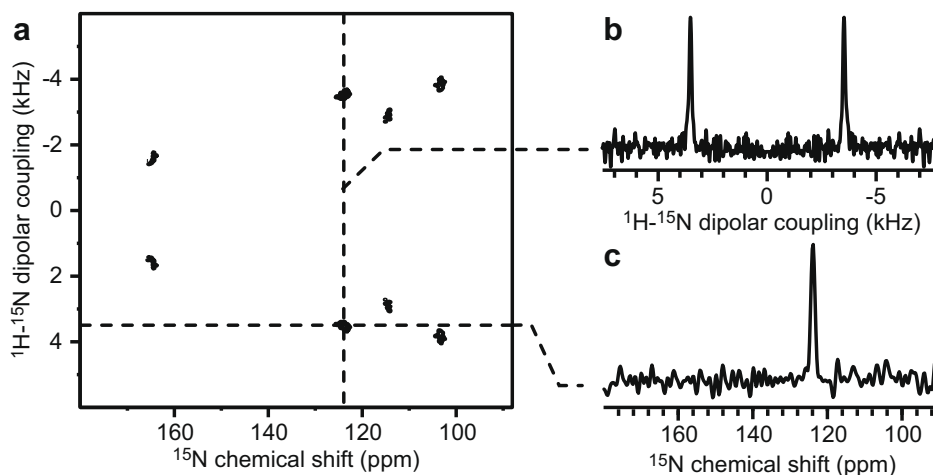
where  $\lambda_e$  is the experimentally determined linewidth in the spectrum processed without apodization.

### 3.5. Experimental case: Two-dimensional $^1\text{H}$ - $^{15}\text{N}$ SLF spectra of a single crystal of NAL

To illustrate the use of controlling the acquisition parameters as described above, we use the acquisition of high-resolution 2D  $^1\text{H}$ - $^{15}\text{N}$  separated-local field (SLF) experiments [10] of a single crystal of NAL as an example. Such a sample serves as a model system for oriented membrane protein systems, which represent a challenging case for NMR due to the generally low sensitivity caused by the low protein density in presence of abundant amounts of lipids. The 2D spectrum of the NAL crystal with an arbitrary orientation acquired using the Polarization Inversion Spin Exchange at the Magic Angle (PISEMA) pulse sequence [11] is shown in Fig. 8. The spectrum contains the expected four resonances, one from each of the four molecules in the unit cell, respectively [12]. In the following, we will investigate the peak located at  $\delta(^{15}\text{N}) = 124$  ppm with a dipolar splitting of  $\sim 7.0$  kHz corresponding to an effective  $^1\text{H}$ - $^{15}\text{N}$  dipole-dipole coupling of  $\sim 8.54$  kHz when encountering the theoretical scaling factor of  $\sqrt{2/3}$  of the SEMA block.

A trace along the  $^{15}\text{N}$  chemical shift dimension (Fig. 8c) of the 2D spectrum sampled with sufficiently large values for  $t_s$  to avoid truncation artefacts in both dimensions ( $t_s(^{15}\text{N}) = 25$  ms,  $t_s(^1\text{H}-^{15}\text{N}) = 10$  ms) reveals a natural linewidth of  $\lambda_n = 59$  Hz and a Gaussian lineshape, while a natural linewidth of  $\lambda_n = 140$  Hz is observed for the indirect  $^1\text{H}$ - $^{15}\text{N}$  dipole-dipole coupling dimension. Such long sampling times result in quite unfavorable sensitivity of the spectrum. For the direct dimension, the sampling time of  $t_s = 25$  ms is negligible compared to the repetition delay of 4 s implying that the overall experiment time is not, to a good approximation, influenced by our choice of  $t_s$ . Hence, we should optimize the SNR for this dimension, while the sensitivity makes no sense. Since  $t_s$  has already been decided, this analysis corresponds to box 1 in Fig. 5. Table 3 reports the normalized SNR ( $n\text{SNR} = 0.671$ ) for the direct dimension indicating that the SNR may be improved by almost 50% by applying matched filtering (Gaussian apodization with  $\lambda_a = \lambda_n$ ) by which  $n\text{SNR} = 1$  is achieved at the expense of having a target linewidth of  $\lambda_t = 83.4$  Hz (see Table 3). Traces of the spectra corresponding to these different parameters are shown in Fig. 9.

While a long sampling time and matched filtering provides the best SNR, it may be desirable to shorten the sampling time in order to reduce the time with high-power decoupling turned on, especially for heat-sensitive samples like membrane proteins. If we



**Fig. 8.** (a) Experimental  $^1\text{H}$ - $^{15}\text{N}$  separated-local field spectrum of a single crystal of NAL. (b) Trace in the indirect heteronuclear dipolar coupling dimension at a chemical shift of 124 ppm. (c) Trace in the direct chemical shift dimension at a dipolar frequency of  $-3.5$  kHz.

**Table 3**  
Experimental parameters for the 2D separated-local field spectrum in Fig. 9.

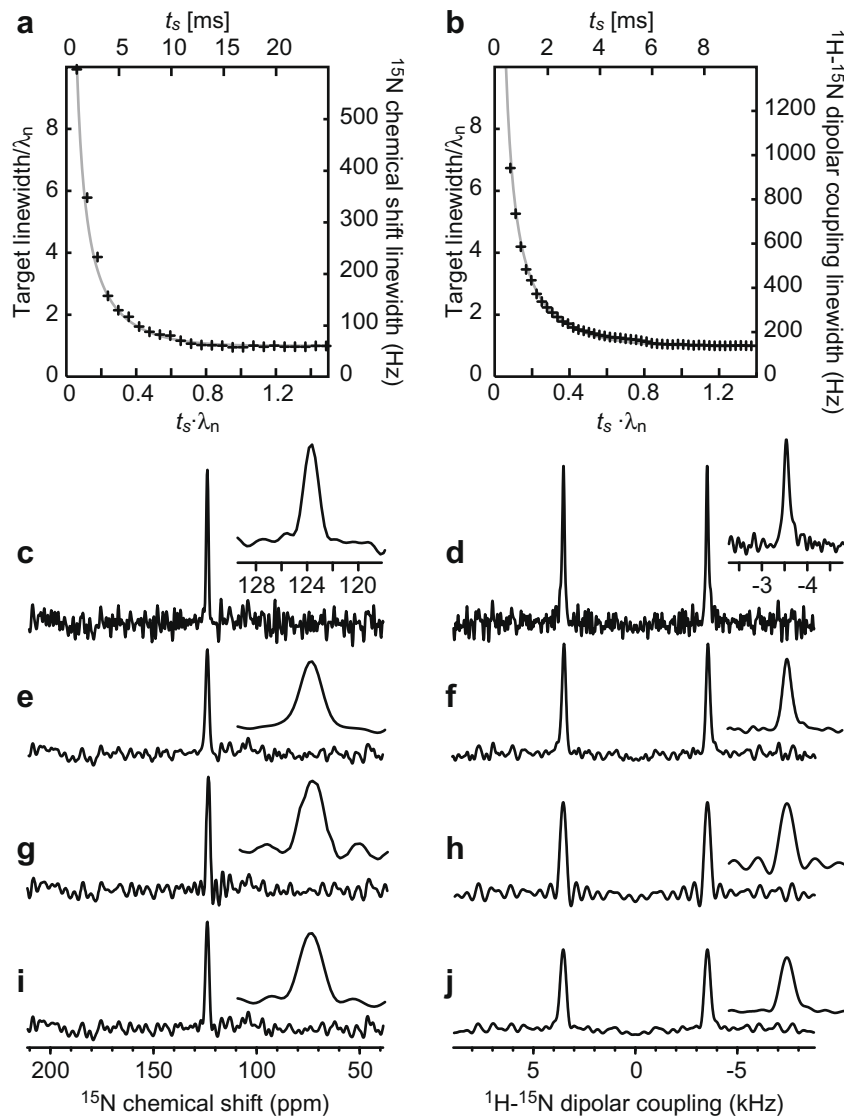
$\lambda_n$ (Hz)	$t_s$ (ms)	np <sup>a</sup>	$\lambda_a$ (Hz)	$\lambda_t$ (Hz)	w	nSNR <sup>b</sup>	sensitivity	Figure index <sup>d</sup>
<sup>15</sup> N chemical shift (direct dimension)								
59	25	1008	0	59	0	0.671	n/a <sup>c</sup>	c
	25	1008	59	83.4	0	1	n/a <sup>c</sup>	e
	8.23	332	0	83.4	0.14	0.941	n/a <sup>c</sup>	g
	10	403	50	83.4	0.05	0.985	n/a <sup>c</sup>	i
<sup>1</sup> H- <sup>15</sup> N dipole-dipole coupling (indirect dimension)								
140	10	250	0	140	0	0.689	0.335	d
	10	250	140	198	0	1	0.487	f
	2	50	0	315	0.21	0.839	0.914	h
	2.68	67	257	315	0.05	0.887	0.835	j

<sup>a</sup> np indicates the number of data points assuming spectral widths of 40,323 Hz and 25,000 Hz for the direct and indirect dimensions, respectively.

<sup>b</sup> Normalized SNR defined in Eq. (15).

<sup>c</sup> The sensitivity is not a relevant parameter for the indirect dimension with such a short  $t_s$  value.

<sup>d</sup> Figure index in Fig. 9.



**Fig. 9.** Illustration of the consequence of varying the acquisition ( $t_s$ ) and processing ( $\lambda_a$ ) parameters for the direct (a, c, e, g, and i) and indirect (b, d, f, h, and j) dimensions of a 2D separated-local field experiment (c.f. Fig. 7). (a and b) Plots of the target linewidth as a function of  $t_s$ . The rightmost vertical and top horizontal scales represent the actual values, while the leftmost vertical and bottom horizontal scales have been normalized with respect to the natural linewidths of  $\lambda_n = 59$  Hz (a) and  $\lambda_n = 140$  Hz (b). The solid grey lines in (a and b) represent the theoretical line achieved by solving Eq. (9). (c, e, g, and i) Traces in the direct dimension at the position indicated in Fig. 7 using the following parameters as also listed in Table 3 ( $t_s$ ;  $\lambda_a$ ): (c) 25 ms; 0, (e) 25 ms; 59 Hz, (g) 8.23 ms; 0, (i) 10 ms; 50 Hz. In (e, g, and i) the target linewidth is 83 Hz. (d, f, and h) Traces in the indirect dimension at the position indicated in Fig. 7 using the following parameters ( $t_s$ ;  $\lambda_a$ ;  $\lambda_t$ ): (d) 10 ms; 0; 140 Hz, (f) 10 ms; 140 Hz; 298 Hz (h) 2 ms; 0; 315 Hz, (j) 2.68 ms; 257 Hz; 315 Hz.

stick to  $\lambda_t = 83.4$  Hz, we may use Eq. (16) to determine that  $t_s = 8.23$  ms is the shortest possible sampling time fulfilling this, using the approach leading to box 3 in Fig. 5. The spectrum, using this sampling time and processed without apodization, is shown in Fig. 9g. The spectrum reveals the expected target linewidth and sacrifices the SNR by only 5.9%, although it also shows clearly visible truncation wiggles, which according to Eq. (11) amount to 14% of the total intensity (see Table 3). If we expect these wiggles to cause problems with the analysis of the spectrum, we may restrict the wiggle height to a smaller value (procedure leading to box 2 in Fig. 5) by increasing  $t_s$  and applying apodization. For example, let us use the restriction  $w = 0.05$ . In this case the solution to Eqs. (16) and (17), using the procedure described in relation to Eq. (20), report that the acquisition and processing should be carried out with the following parameters:  $t_s = 10$  ms,  $\lambda_a = 50$  Hz (Fig. 9i). This leads to the same target linewidth as before but with a reduced wiggle height, and a SNR, which is only 1.5% lower than the maximum achievable (Table 3).

Addressing the indirect dimension, we are now in a case where the overall experiment time is directly proportional to  $t_s$ , provided we maintain a fixed value for the spectral width. In this case, the focus will be to optimize the *sensitivity*, since this parameter reports how efficiently we accumulate the data in the indirect dimension. Clearly, the set of parameters used for the spectrum in Figs. 8 and 9d provide a high resolution but a very poor sensitivity of 0.335 (Table 3). Applying matched filtering (Fig. 9f) improves this number somewhat (0.487), but it is obvious that matched filtering is not the means to optimize the *sensitivity*. For large membrane protein samples [13], it will typically not be possible to record that many  $t_1$  increments, and hence  $t_s$  values on the order of a few ms are commonly applied [14–17]. An analysis of employing a sampling time of  $t_s = 2$  ms (procedure leading to box 1 in Fig. 5) shows that the target linewidth will be  $\lambda_t = 315$  Hz (Table 3, Fig. 9h) and that the spectrum will display substantial wiggles, amounting to 21% of the peak height. However, the *sensitivity* is roughly doubled (0.914) compared to what is achieved using the long sampling time of 10 ms. Here, the intense wiggles may be problematic, so it may be desirable to reduce the wiggle height using the approach described in box 2 in Fig. 5. Using the parameters  $w = 0.05$  and  $\lambda_t = 315$  Hz as an example, we find that this may be achieved using  $t_s = 2.68$  ms and  $\lambda_a = 257$  Hz with a *sensitivity* of 0.835 (Table 3, Fig. 9j). Immediately, one would believe that the increase of  $t_s$  from 2 to 2.68 ms results in a 1.34-fold increase in the overall experiment time, but as seen from the ratios of the sensitivities  $0.835/0.914 = 0.91$ , the latter experiment will only lack 9% of the intensity if the two experiments were acquired using the same overall experiment time.

In the discussion above, we have not dealt with cases where the natural linewidth varies between peaks in the spectrum. Clearly, this will often be the case, especially for solid samples where, e.g., the efficiency of the decoupling may vary with the nuclear spin interactions leading to different values for  $\lambda_n$  for different resonances. Such variations in natural linewidth will obviously also be reflected in the target linewidth, but for Gaussian resonances, which are most typical for solid samples, the effect will be reduced due to the square root dependence (see Eq. (22)). To give an idea on the variations in the target linewidth and wiggle height caused by variations in  $\lambda_n$ , we consider the direct dimension of the 2D SLF experiment, sampled using  $t_s = 10$  ms (Fig. 9i). Variations in the natural linewidth ( $\lambda_n = 59$  Hz) by  $\pm 20\%$  will lead to variations in  $\lambda_t$  by less than 10% and changes in the wiggle height smaller than 0.03. Consequently, estimation of an approximate average value for the natural linewidth will be sufficient to provide a fair prediction of the spectral attributes. In cases where the natural linewidth is completely unknown, we may still estimate the wiggle height for

each peak using the plot in Fig. 7 once the target linewidth is measured.

#### 4. Experimental

The NAL crystals were produced as described previously [18]. For the solid-state NMR experiments, a high-quality crystal of approximate size  $0.5 \times 0.5 \times 0.3$  mm<sup>3</sup> was selected. The experiment was performed on a Bruker Avance 400 spectrometer with a 9.4 T magnet with a <sup>1</sup>H and <sup>15</sup>N resonance frequencies of 400.13 and 40.55 MHz employing a 4 mm Bruker triple-resonance MAS probe in double-resonance <sup>1</sup>H–<sup>15</sup>N configuration. The PISEMA [11] pulse sequence employed a <sup>1</sup>H RF field strength of 100 kHz for the pulses and decoupling period, which used SPINAL-64 decoupling [19], while it was reduced to 41.8 kHz during the 1 ms CP period as well as the  $t_1$  period. The <sup>15</sup>N RF field strength was 41.8 kHz during CP and 50 kHz during  $t_1$  to meet the Hartmann–Hahn condition with the off-resonance field on <sup>1</sup>H. The  $t_1$  dimension employed 250 increments of 40  $\mu$ s corresponding to a spectral width ( $sw$ ) of 25 kHz and  $t_s = 10$  ms, while the direct dimension was sampled using  $sw = 40,323$  Hz and  $t_s = 25$  ms. Each  $t_1$  value was acquired using 8 scans with a repetition delay of 4 s.

To evaluate the consequences of different  $t_s$  values, as investigated in Fig. 9, only the appropriate number of points, calculated as  $np = sw \cdot t_s$  (listed in Table 3), were used and zero-filled to 2048/4096 points in the indirect/direct dimensions. All spectra were processed using SIMPSON [20–22].

#### 5. Conclusion

We have presented a thorough analytical and numerical analysis of the consequences of appropriate definition of the dimensions of the sampling space and the associated use of apodization to improve spectral quality. This leads to simple procedures to setup optimal sampling used to define the points inbetween time 0 and the maximum sampling time derived here. We anticipate that our simple protocol will find widespread application for optimizing the outcome of multiple-dimensional NMR experiments where the length of the experiments and potential sample heating effects through decoupling during acquisition may be important issues.

#### Acknowledgments

We acknowledge support from the Danish National Research Foundation, the Danish National Advanced Technology Foundation, the Danish Natural Science Research Foundation, the Danish Center for Scientific Computing, and Carlsbergfondet.

#### References

- [1] G. Bodenhausen, H. Kogler, R.R. Ernst, Selection of coherence transfer pathways in NMR experiments, *J. Magn. Reson.* 58 (1984) 370–388.
- [2] M.H. Levitt, P.K. Madhu, C.E. Hughes, Cogwheel phase cycling, *J. Magn. Reson.* 155 (2002) 300–306.
- [3] H. Heise, K. Seidel, M. Etkorn, S. Becker, M. Baldus, 3D NMR spectroscopy for resonance assignment and structure elucidation of proteins under MAS: novel pulse schemes and sensitivity considerations, *J. Magn. Reson.* 173 (2005) 64–74.
- [4] K. Aggarwal, M.A. Delsuc, Triangular sampling of multidimensional NMR data sets, *Magn. Reson. Chem.* 35 (1997) 593–596.
- [5] J.C. Hoch, A.S. Stern, *NMR Data Processing*, John Wiley & Sons, Hoboken, New Jersey, 1996.
- [6] A.S. Stern, K.B. Li, J.C. Hoch, Modern spectrum analysis in multidimensional NMR spectroscopy: comparison of linear-prediction extrapolation and maximum-entropy reconstruction, *J. Am. Chem. Soc.* 124 (2002) 1982–1993.
- [7] D. Rovnyak, D.P. Frueh, M. Sastry, Z.Y. Sun, A.S. Stern, J.C. Hoch, G. Wagner, Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction, *J. Magn. Reson.* 170 (2004) 15–21.

- [8] R.R. Ernst, G. Bodenhausen, A. Wokaun, Principles of Nuclear Magnetic Resonance in One and Two Dimensions, Clarendon Press, Oxford, 1987.
- [9] Mupad version 4.0. Homepage. Available from: <http://www.sciface.com>.
- [10] R.K. Hester, J.L. Ackerman, B.L. Neff, J.S. Waugh, Separated local field spectra in NMR: determination of structure of solids, Phys. Rev. Lett. 36 (1976) 1081.
- [11] C.H. Wu, A. Ramamoorthy, S.J. Opella, High-resolution heteronuclear dipolar solid-state NMR spectroscopy, J. Magn. Reson. A 109 (1994) 270–272.
- [12] P.J. Carroll, P.L. Stewart, S.J. Opella, Structures of two model peptides: *N*-acetyl-d,l-valine and *N*-acetyl-l-valyl-l-leucine, Acta Cryst. C 46 (1990) 243–246.
- [13] T. Vosegaard, N.C. Nielsen, Towards high-resolution solid-state NMR on large uniformly  $^{15}\text{N}$ - and [ $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labeled membrane proteins in oriented lipid bilayers, J. Biomol. NMR 22 (2002) 225–247.
- [14] M. Kamihira, T. Vosegaard, A.J. Mason, S.K. Straus, N.C. Nielsen, A. Watts, Structural and orientational constraints of bacteriorhodopsin in purple membranes determined by oriented-sample solid-state NMR spectroscopy, J. Struct. Biol. 149 (2005) 7–16.
- [15] S.H. Park, S. Prytulla, A.A. De Angelis, J.M. Brown, H. Kiefer, S.J. Opella, High-resolution NMR spectroscopy of a GPCR in aligned bicelles, J. Am. Chem. Soc. 128 (2006) 7402–7403.
- [16] T. Vosegaard, M. Kamihira-Ishijima, A. Watts, N.C. Nielsen, Helix conformations in 7TM membrane proteins determined using oriented-sample solid-state NMR with multiple residuespecific  $^{15}\text{N}$  labeling, Biophys. J. 94 (2008) 241–250.
- [17] R. Mahalakshmi, F.M. Marassi, Orientation of the *Escherichia coli* outer membrane protein ompX in phospholipid bilayer membranes determined by solid-state NMR, Biochemistry 47 (2008) 6531–6538.
- [18] K. Bertelsen, J.M. Pedersen, N.C. Nielsen, T. Vosegaard, 2D separated-local field spectra from projections of 1D experiments, J. Magn. Reson. 184 (2007) 273–279.
- [19] B.M. Fung, A.K. Khitrin, K. Ermolaev, An improved broadband decoupling sequence for liquid crystals and solids, J. Magn. Reson. 142 (2000) 97–101.
- [20] M. Bak, J.T. Rasmussen, N.C. Nielsen, SIMPSON: a general simulation program for solid-state NMR spectroscopy, J. Magn. Reson. 147 (2000) 296–330.
- [21] T. Vosegaard, A. Malmendal, N.C. Nielsen, The flexibility of SIMPSON and SIMMOL for numerical simulations in solid- and liquid-state NMR spectroscopy, Monatsh. Chem. 133 (2002) 1555–1574.
- [22] Z. Tosner, T. Vosegaard, C. Kehlet, N. Khaneja, S.J. Glaser, N.C. Nielsen, Optimal control in NMR spectroscopy: numerical implementation in SIMPSON, J. Magn. Reson. 197 (2009) 120–134.