JOURNAL OF MAGNETIC RESONANCE **134,** 184–188 (1998)
ARTICLE NO. MN981500

Influence of Noise on Peak Integrals Obtained by Direct Summation

S. Bourg and J. M. Nuzillard[1]

Laboratoire de Pharmacognosie, UPRESA 6013, CPCBAI, Bat. 18 Moulin de la Housse, 51097 Reims Cedex 2, France

Received December 11, 1997; revised May 6, 1998

The influence of zero-filling on the peak integral precision is examined when integration is carried out by direct summation of spectral point ordinates. A relationship that allows the computation of the standard deviation of the integrals is derived, taking into account the effects of apodization and noise intensity profile. Discrepancies between theory and experimentation arise from nonideal characteristics of noise. © 1998 Academic Press

Key Words: **integration; zero-filling; apodization.**

Quantitative information in NMR spectra is brought by peak areas (*1, 2*). Two methods of peak integration are used: direct summation of spectral point ordinates, and peak parameter search by curve fitting. In the absence of a model for the peak shape, direct summation is the only practical technique. It is not, however, adapted to partially overlapping peaks. The precision of both techniques has been investigated, taking into account the possible sources of systematic and random errors (*2–6*). The aim of this communication is to provide a practical relationship that evaluates the random error on data points sums as a function of the spectral noise level, the number of acquired data points, the shape of an eventual apodization function, the level of zero-filling, and the width of the integration region. Our initial motivation for this work was the quantitative analysis of polymeric reaction mixtures by ^{13}C NMR spectroscopy, where chemical shift diversity for the same chemical functionality and low signal-to-noise ratios make peak modeling very unpractical. The result of this study, however, is of general applicability.

Integration by direct summation introduces two kinds of systematic errors. One is due to the approximation caused by the assimilation of the integral of a continuous function with a finite sum (*4*); the other one is caused by the parts of the peaks that are left outside of the integration range (*6*). Methods can be found to minimize these errors, but choosing an integration interval requires some prior knowledge of peak width and shape. Clearly, the problem caused by the digitization of the frequency axis is at best solved by zero-filling. This procedure does not contribute any information, but performs data interpolation (*7*). If neither apodization nor zero-filling is applied to an FID, the spectral noise is uncorrelated (providing that spectrometer noise is uncorrelated as well). The variance of the sum of data points is then the sum of the variances of the individual spectral points. It will be shown that the statistical behavior of a point sum becomes more complex even when a single level of zero-filling is used.

Consider an FID made of n complex points (indexed by k), completed to N complex points by zero-filling, and its Fourier transform. The spectrum is phased and its imaginary part discarded. The first (for commodity) m points are summed up (indexed by l). In order to establish statistics, the whole experiment is repeated M times (indexed by j). The procedure is very similar to a Monte Carlo parameter variance estimation (*8*). Let x_{jk} and $\epsilon_{jk} = \epsilon'_{jk} + i\epsilon''_{jk}$ be the FID value and its noise at time t_k for the jth experiment. The noise has a zero mean, and its standard deviation σ_k at time t_k is defined by

$$\sigma_k^2 = \frac{1}{M}\sum_{j=1}^{M}(\epsilon'_{jk})^2 = \frac{1}{M}\sum_{j=1}^{M}(\epsilon''_{jk})^2. \quad [1]$$

It will be assumed first that σ_k does not vary with k. The global FID noise σ_F is defined by

$$\sigma_F^2 = \frac{1}{n}\sum_{k=0}^{n-1}\sigma_k^2. \quad [2]$$

The sum of the m first points of the spectrum can be written (*9*)

$$S_j = \sum_{l=0}^{m-1}\sum_{k=0}^{n-1} x_{jk}\exp(-2i\pi kl/N). \quad [3]$$

The error s_j introduced by the noise on the real part of S_j is then

$$s_j = \sum_{k=0}^{n-1}\left(\epsilon'_{jk}\cdot\sum_{l=0}^{m-1}\cos(2\pi kl/N) + \epsilon''_{jk}\cdot\sum_{l=0}^{m-1}\sin(2\pi kl/N)\right). \quad [4]$$

[1] To whom correspondence should be addressed.

1090-7807/98 $25.00

Its variance $\sigma_I^2 = (1/M)(\Sigma_{j=1}^{M} s_j^2)$ can be evaluated as

$$\sigma_I^2 = \sigma_F^2 \sum_{k=0}^{n-1}((\sum_{l=0}^{m-1} \cos(2\pi kl/N))^2 + (\sum_{l=0}^{m-1} \sin(2\pi kl/N))^2). \quad [5]$$

In the development of s_j^2, terms such as $\epsilon'_{jk}\,\epsilon'_{jk'}$ $(k \neq k')$, $\epsilon''_{jk}\,\epsilon''_{jk'}$, or $\epsilon'_{jk}\,\epsilon''_{jk'}$ are discarded, as their average value must be zero if the noise samples are statistically independent. The signal x no longer intervenes, proving that the error does not depend on the shape of the integrated signal. Terms in Eq. [5] will be developed in two ways. First, by expanding the squares of the sums over l, it will be possible to analytically compare the variance of the S_j when either no or a single zero-filling is performed:

$$\sigma_I^2 = \sigma_F^2 \sum_{k=0}^{n-1}(\sum_{l=0}^{m-1} 1 + 2\sum_{l=0}^{m-1}\sum_{l'=l+1}^{m-1} \cos(2\pi k(l-l')/N)). \quad [6]$$

Defining the index l'' by $m - l'' = l' - l$ makes

$$\sigma_I^2 = \sigma_F^2(mn + 2\sum_{l''=1}^{m-1} l'' \sum_{k=0}^{n-1} \cos(2\pi k(m-l'')/N)). \quad [7]$$

If $N = n$ (no zero-filling), the sum over k is always zero and therefore $\sigma_I^2 = mN\sigma_F^2$. This result is the expected one: the variance of the spectral noise is $N\sigma_F^2$, and m spectral points are added together. The limiting case where $m = N$ leads to $\sigma_I = n\sigma_F$ as expected, because $S_j = Nx_{j0}$. This last result is true whatever the zero-filling level.

If $N = 2n$, the summation over k in Eq. [7] yields either 0 or 1, depending on the parity of $m - l''$. Considering even values of m leads to

$$\sum_{l''=1}^{m-1} l'' \sum_{k=0}^{n-1} \cos(2\pi k(m-l'')/N) = \left(\frac{m}{2}\right)^2 \quad [8]$$

as being the sum of the $m/2$ first odd numbers. Therefore,

$$\sigma_I^2 = \sigma_F^2(mn + m^2/2) \ (m \text{ even})$$

$$\sigma_I^2 = \sigma_F^2(mn + (m^2-1)/2) \ (m \text{ odd, not proved}). \quad [9]$$

Both expressions are practically identical when m is bigger than a few units. The spectral noise variance is still $n\sigma_F^2$ (corresponding to the case $m = 1$); there is no improvement of the spectral signal-to-noise ratio upon zero-filling (*10*).

Let m_1 be the number of points in the integration range in absence of zero-filling, and z the ratio N/n that defines the extent of the zero-filling. In order to preserve the integration range in frequency units, m must be proportional to m_1: $m = zm_1$. We then get

$$\sigma_I = z\sigma_F(nm_1)^{1/2} \text{ when } z = 1 \quad [10a]$$

$$\sigma_I = z\sigma_F(1/2)^{1/2}(nm_1)^{1/2} \text{ when } z = 2 \text{ and } m_1 \ll n. \quad [10b]$$

Considering for the moment that the digitization error does not intervene, the value of the integral is proportional to z, like σ_I, and therefore the *integral-to-noise ratio* is improved by a factor of $\sqrt{2}$ when one level of zero-filling is used and if the width of the integration zone is much smaller than the spectral width. Even though the spectral noise in both $z = 1$ and $z = 2$ cases is said to be noncorrelated, its statistical behavior is not the same (*6*). The noise integral variance may vary linearly ($z = 1$) or quadratically ($z = 2$) as a function of m.

In order to investigate the role of higher z values, Eq. [5] can be rewritten

$$\sigma_I^2 = \sigma_F^2(\sum_{k=0}^{n-1}|\sum_{l=0}^{m-1} \exp(2i\pi kl/N)|^2) \quad [11]$$

or

$$\sigma_I^2 = \sigma_F^2 \sum_{k=0}^{n-1}\left(\frac{\sin(km\pi/N)}{\sin(k\pi/N)}\right)^2. \quad [12]$$

Plots of σ_I^2 as a function of m show that for $m \geq z$ (or $m_1 \geq 1$!), σ_I^2 can be very well approximated by

$$\sigma_I^2 \approx \frac{1}{2}(z\sigma_F)^2 m_1 n\left(1 + \frac{m_1}{n}\right) \quad [13]$$

or even

$$\sigma_I^2 \approx \frac{1}{2}(z\sigma_F)^2 m_1 n \quad [14]$$

if the integration range is much smaller than the spectral width, which is usually the case. The expression established rigorously for $z = 2$ is still valid for higher zero-filling levels. There is virtually no point in increasing z beyond 2, except in the hope of reducing the systematic error due to the digitization of the frequency axis.

The processing of an FID generally includes an apodization step in order to both optimize the peak height to spectral noise ratio and to limit the peak extension due to FID truncation (*11*). Let a_k be the value of the apodization function at time t_k. Equation [12] becomes

$$\sigma_I^2 = \sigma_F^2 \sum_{k=0}^{n-1}\left(a_k \frac{\sigma_k}{\sigma_F}\frac{\sin(km\pi/N)}{\sin(k\pi/N)}\right)^2 \quad [15]$$

when both apodization and FID noise distribution are taken

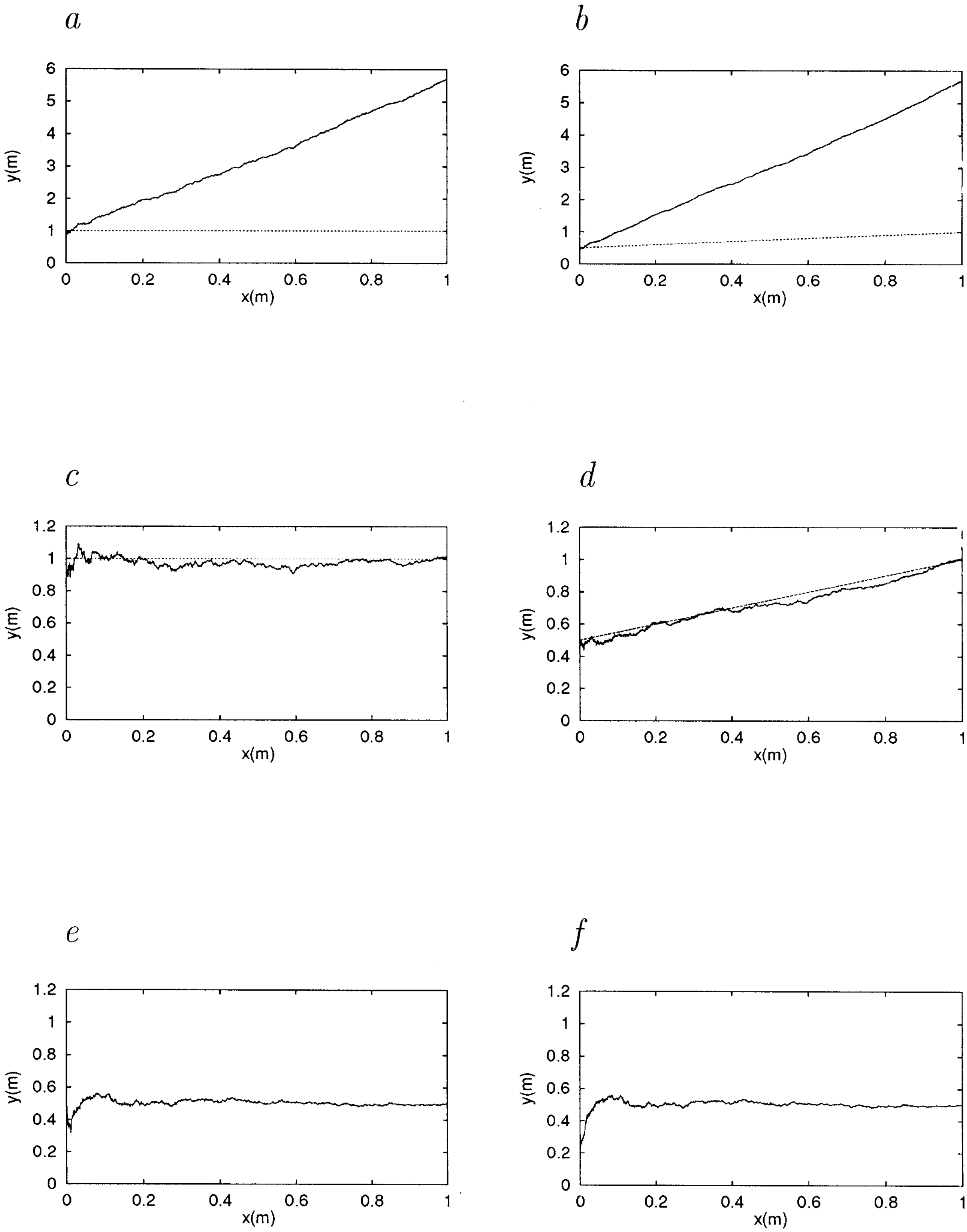

FIG. 1. Plots of the reduced variance of the noise integral versus the reduced integration range, for signals filtered through an analog filter. Thin dotted lines correspond to the theoretical values, computed from Eq. [20]. No apodization was performed. (a) No zero-filling was applied; (b) one level of zero-filling was applied ($z = 2$). Graphs (c) and (d) are identical to (a) and (b), but with the first FID point multipled by 0.42 (see text). (e) No zero-filling; (f) one level of zero-filling, for digitally filtered signals.

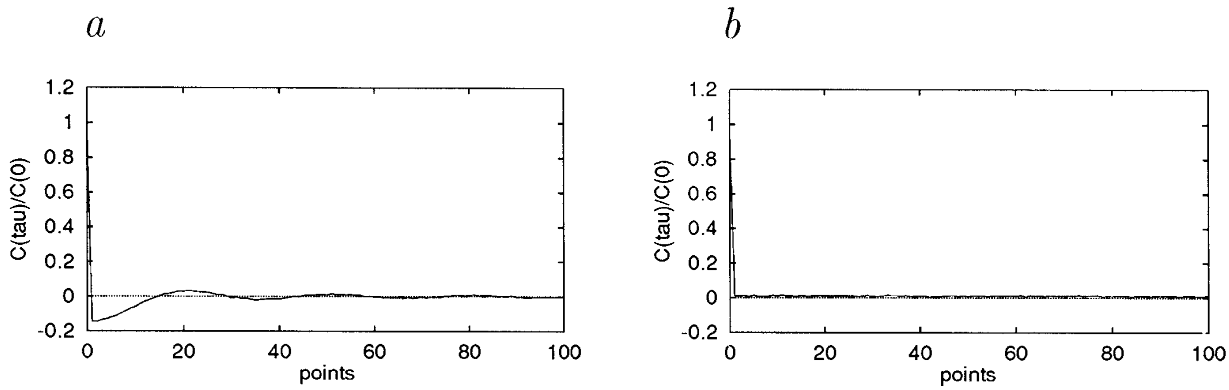

FIG. 2. Normalized autocorrelation functions $C(\tau)$ of the spectrometer noise. They are computed by Fourier transformation of the power spectra of the FID, and averaged over the 512 data records (*9*). Only the real part is shown. Normalization consists of dividing $C(\tau)$ by $C(0)$ for comparison purposes. (a) Digital filter: the bandwidth was 20 kHz and the sampling rate was 10 kHz. (b) Analog filter: in the same conditions as in (a).

into account. The values $\rho_k = \sigma_k/\sigma_F$ define the noise profile of the FID. It is an instrumental characteristic that should be as close as possible to the unit function. The noise parameter that is experimentally accessible is generally not the FID noise level, especially when the FID is truncated. Conversely, there is generally a zone in the spectrum where the noise level σ_S can be measured. The relationship between σ_F and σ_S is given by

$$\sigma_S^2 = \sigma_F^2 \sum_{k=0}^{n-1} (a_k \rho_k)^2, \qquad [16]$$

providing that the FID noise is not time correlated. The general relationship that allows to evaluate the integral noise level is then

$$\sigma_I^2 = \sigma_S^2 \sum_{k=0}^{n-1} \left(a_k \rho_k \frac{\sin(km\pi/N)}{\sin(k\pi/N)} \right)^2 \Big/ \sum_{k=0}^{n-1} (a_k \rho_k)^2. \qquad [17]$$

The denominator in Eq. [17] can be computed for the widely used Lorentzian and Gaussian apodization functions, if the FID noise profile is flat and if the apodization function is not truncated ($a_{n-1} \approx 0$). A line broadening $\Delta\nu$ and an acquisition time T give

$$\sum_{k=0}^{n-1} a_k^2 = \left(\frac{1}{2\pi}\right) \frac{n}{T\Delta\nu} \text{ for a Lorentzian apodization and} \qquad [18a]$$

$$\sum_{k=0}^{n-1} a_k^2 = \left(\frac{\ln 2}{2\pi}\right)^{1/2} \frac{n}{T\Delta\nu} \text{ for a Gaussian apodization.} \qquad [18b]$$

An experimental validation of Eq. [13] was achieved by recording $M = 512$ FIDs of $n = 1024$ complex noise data points. As previously mentioned, the shape of the integrated signal does not intervene. Integration over a spectral region that contains either a signal or only noise results in identical random variations. Therefore, no pulse was sent to the sample and no relaxation delay was needed. Four scans per FID with receiver phases x, y, $-x$, $-y$ were used in order to compensate for eventual demodulation DC offset and channel balance defaults. A spectral width of 10 kHz was chosen. The analog filter bandwidth was set to 20 kHz, and the digital filter was disabled. Zero-fillings by factors 1, 2, 4, 8, and 16 were applied. After Fourier transformation, the spectral noise and the sums of the first m real points were calculated for each experiment. Variances of the noise integrals were then determined. The Eq. [13] can be rewritten as

$$\frac{1}{mz}\left(\frac{\sigma_I}{\sigma_S}\right)^2 = \frac{1}{2}\left(1 + \frac{m}{N}\right) \quad (z \geq 2). \qquad [20]$$

Plots of the left-hand side expression $y(m)$ as a function $x(m) = m/N$ are drawn in Figs. 1a and 1b. The functions $x(m)$ and $y(m)$ are the reduced integration range and the reduced variance of noise integral, respectively. The values of z greater or equal to 2 give identical graphs for $m \geq z$. When $z = 1$, Eq. [10a] becomes $y(m) = 1$, independently of m. The theoretical graphs are clearly different from the experimental ones, even though the discrepancy is small when $x(m) \ll 1$. A closer look at the acquired data matrix shows that at any time t_k, the variance of the noise is independent of k, except at time t_0. At this time, σ_0^2 is 5.6 times higher than at the other times, both in the real and the imaginary part of the noise. This is rather coherent with $y(N) = 5.6$ instead of 1 (see Figs. 1a and 1b), whatever the applied zero-filling. From Eq. [17],

$$y(N) = n(a_0\rho_0)^2 \Big/ \sum_{k=0}^{n-1} (a_k \rho_k)^2, \qquad [21]$$

which reduces to $y(N) \approx \rho_0^2$ in the present case. Therefore, a good way of matching theory and experiment is to multiply the first FID point by $y^{-1/2}(N)$, 0.42 in our case. The result is presented in Figs. 1c and 1d. It is interesting to note that the factor 0.42 is close to the one, 0.5, that is used to position the spectral baseline at the zero level (*1*). It can be shown by means of Eq. [17] that multiplying the first data point by a correction factor does not influence the variance of the integral if the integration range is narrow compared to the spectral width.

Another experiment was performed using identical conditions, but with the digital filtering switched on. The noise that comes out of the digital filter is clearly not uniformly distributed. The noise profile was determined and used in Eq. [17]. It was not possible to interpret the experimental plots of $y(m)$ versus $x(m)$ drawn in Figs. 1e and 1f. A possible correlation of the spectral noise was tested. The autocorrelation function of the spectral noise (*9*) is drawn in Fig. 2a. The ideal profile produced by the analog filter is not obtained (see Fig. 2b). However, plots show that $y(m) \approx 0.5$ for all m and z values. This experimental result is useful for the measurement of integral precision, even though it is not fully supported by theory.

In conclusion, determining uncertainties on peak integrals obtained by direct summation requires the prior knowledge of the instrumental noise characteristics. Even though zero-filling does not participate in the improvement of the spectral signal-to-noise ratio, it may increase the integral precision by a factor up to $2^{1/2}$ when the time-domain noise is not correlated. A rough evaluation of the integral variance may be given by $\sigma_I^2 = mz\sigma_S^2/2$, providing that zero-filling was used. The experimental approach used for digitally filtered signals should be a prerequisite to any quantitative spectral analysis by means of direct peak integration.

ACKNOWLEDGMENT

SB thanks the Département de la Marne (France) for financial support.

REFERENCES

1. J. Led and H. Gesmar, *Methods Enzymol.* **239,** 318 (1994).
2. Y.-L. Martin, *J. Magn. Reson. A* **111,** 1 (1994).
3. D. W. Posener, *J. Magn. Reson.* **14,** 121 (1974).
4. G. H. Weiss, J. E. Kiefer, and J. A. Ferretti, *Chemom. Intell. Lab. Sys.* **4,** 223 (1988).
5. L. Chen, C. E. Cottrell, and A. G. Marshall, *Chemom. Intell. Lab. Sys.* **1,** 51 (1986).
6. C. Rischel, *J. Magn. Reson. A* **116,** 255 (1995).
7. Z. Zolnai, N. Juranic, J. L. Markley, and S. Macura, *J. Magn. Reson. A* **119,** 53 (1996).
8. W. H. Press, B. P. Flannery, S. A. Teutolsky, and W. T. Vetterling, “Numerical Recipes in C,” 2nd ed., Cambridge University Press, Cambridge (1992).
9. A. G. Marshall and F. R. Verdun, “Fourier Transforms in NMR, Optical, and Mass Spectrometry,” Elsevier, Amsterdam (1990).
10. R. Freeman, “A Dictionary of NMR Concepts,” Longman, Harlow, UK (1988).
11. R. R. Ernst, *in* “Advances in Magnetic Resonance” (J. S. Waugh, Ed.), Vol. 2, p. 1, Academic Press, New York (1966).