Post-processing of individual signals for de-noising

Jun Fukazawa, Kazuyuki Takeda, K. Takegoshi *

Department of Chemistry, Graduate School of Science, Kyoto University, Kitashirakawa Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan

ARTICLE INFO

ABSTRACT

In this work, we examine the fluctuation of the intensity and the phase of an NMR signal during repetition of experiments and investigate possibilities of using these information to judge suspicious peaks, whose true colors may be noises or genuine signals. We firstly analyze the intensity and the phase of an NMR signal separately, and show that for the accumulated spectral profile the contribution of the intensity is less than that of the phase. Secondly we show that we can de-noise a noisy spectrum by using the standard deviation of phase at each spectral point. We then compare the de-noising effect of the present approach and that of the phase-covariance method proposed recently, which is an alternative method of appreciating phase distribution. Finally, effects of the dispersion component are discussed.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

The sensitivity of nuclear magnetic resonance (NMR) spectroscopy is low. In many practical cases, resonance lines in the spectrum are buried in the noise, when the data is acquired only once. To improve the signal-to-noise ratio (SNR), it is common to accumulate a number of free induction decays (FIDs). When N FIDs are added together, SNR is increased by $\sqrt{N}$ compared to that of the single scan signal. Since this increase rate is rather slow, measurements often require very large N, making experiments quite time consuming. Historically, the conventional signal accumulation scheme has been a natural choice, considering the limited data storage space. On the other hand, recent progress in storage devices has made separate data recording feasible, as demonstrated by Ivchenko et al. in the context of multiplex phase cycling [1].

In this work, we propose to store the individual data separately and statistically analyze the data. We show that such analysis gives us additional information that can be used to process the data in a more efficient manner than mere accumulation. In fact, we recently proposed a signal analysis method for de-noising based on phase correlation between the NMR signal and the excitation pulse (the phase-covariance analysis [2]). In the following we analyze the distribution of the intensity and phase of the data points in terms of variance and covariance, and present their applications to de-noising a noisy spectrum and elimination of spurious signals.

In this paper we propose two new NMR data processing methods, referred to as phase standard deviation weighting (PSDW) and

phase covariance weighting (PCW), and examine these methods using experimental data, and show that PSDW and PCW are useful to judge signals and noises.

2. Theory, results, and discussion

2.1. A model of a quadrature NMR signal and noise

A quadrature NMR FID of a single signal at Ω_0 may be given by

$$s_0(t) = \begin{cases} M_0 \exp\left(-i\Omega_0 t - \frac{t}{T_2}\right), & (t \geqslant 0) \\ 0, & (t < 0) \end{cases}, \quad (1)$$

where we assumed no noise involved and a Lorentzian lineshape with the spin–spin relaxation time of T_2. $s_0(t)$ is Fourier transformed to give the spectrum written as

$$S_0(\Omega) = \frac{1}{\sqrt{2\pi}} \int_0^\infty M_0 \exp\left(-i\Omega_0 t - \frac{t}{T_2}\right) \cdot \exp(i\Omega t) dt$$
$$= M(v) \exp(-i\phi_0(v)), \quad (2)$$

with $v = \Omega - \Omega_0$, where $M(v)$ and $\phi_0(v)$ are written as

$$M(v) = \frac{M_0}{\sqrt{2\pi\left(\frac{1}{T_2^2} + v^2\right)}}, \quad (3)$$

and

$$\phi_0(v) = -\text{Arc}\tan(T_2 v), \quad (4)$$

respectively. At the center of the peak of $S_0(\Omega)(\Omega = \Omega_0)$, the phase factor is zero ($\phi_0(0) = 0$), while it reaches $\pm\pi/2$ at the far envelope.

* Corresponding author. Fax: +81 75 753 4011.
E-mail address: takeyan@kuchem.kyoto-u.ac.jp (K. Takegoshi).

This phase is related to the dispersion (the imaginary) component of the signal, and we shall hereafter refer to this phase as an intrinsic phase.

More generally, a quadrature accumulated NMR signal at a frequency v may be given by the product of an intensity $I(v)$ and a phase factor as

$$S(v) = \overline{I(v) \cdot \exp(-i(\phi(v) + \xi(v)))}, \tag{5}$$

where $\phi(v)$ is the intrinsic phase and $\xi(v)$ represents the phase ascribed to the phase difference between the transmitter and the receiver systems and the frequency-dependent phase shift due to the experimental time delays, etc. Since this phase can be removed by the conventional phase correction, the phase $\xi(v)$ is neglected hereafter for simplicity. The overline in Eq. (5) denotes the averaging over the accumulation;

$$\overline{I(v) \cdot \exp(-i\phi(v))} = \frac{1}{N}\sum_{j=1}^{N}(I_j(v) \cdot \exp(-i\phi_j(v))), \tag{6}$$

where $I_j(v)$ and $\phi_j(v)$ are the intensity and the phase of the signal at v in the j-th spectrum of the N accumulation. Each $I_j(v)$ and $\phi_j(v)$ includes noise as

$$I_j(v) \cdot \exp(-i\phi_j(v)) = I_0(v) \cdot \exp(-i\phi_0(v)) + \delta_I^j(v) \cdot \exp(-i\delta_\phi^j(v)), \tag{7}$$

where $I_0(v)$ and $\phi_0(v)$ are the intensity and the phase without noise, and $\delta_i^j(v)$ and $\delta_\phi^j(v)$ are the intensity and the phase of noise at v in the j-th spectrum, respectively (Fig. 1a).

Fig. 1b describes two cases, $I_0 > \delta_I^j$ and $I_0 \lesssim \delta_I^j$. When the intensity of the signal is larger than the intensity of the noise, the apparent distribution of the phase would be relatively small. For a pure signal without the noise, the standard deviation σ_ϕ of the phase is zero. On the other hand, the phase distribution and σ_ϕ would increase with the noise. At frequencies where $I_0 = 0$, only the noise exists and σ_ϕ becomes $\sqrt{\int_{-\pi}^{\pi} \phi^2 d\phi / 2\pi} = \frac{\pi}{\sqrt{3}}$. Hence, σ_ϕ at each frequency carries information as to certainty about whether the signal exists. It may be also envisaged that the standard deviation of the signal intensity σ_I would not be significantly different for the signal and the noise.

2.2. Distribution of phase and intensity

In order to examine the distribution of the intensity and the phase experimentally, we performed ^{13}C NMR of a low-concentration mixture of D,L-alanine (3 wt.%) and glycine (0.7 wt.%) in KBr powder. The conventional combined techniques of cross-polarization (CP) and decoupling sequence (Fig. 2) under magic-angle spinning (MAS) of the sample was used to observe high-resolution ^{13}C solid state NMR spectra in 14 T with a repetition time of 20 s. We applied the conventional phase alternation for the 1st 90° pulse and the pairs of FIDs were subtracted with each other to reduce artifacts [3]. Total 1800 pairs of FIDs were collected separately and Fourier transformed to obtain an array of spectra. Fig. 3b shows one of them, showing that the present spectrum corresponds to the case of $I_0 \lesssim \delta_I^j$. The accumulated spectrum is shown in Fig. 3a.

Firstly, we examined distributions of the intensity δ_I and phase δ_ϕ of the pure noise. The δ_I values obtained from 100 noise points indicated in Fig. 3a by two horizontal arrows (i.e. 1800×100 data points) are plotted in Fig. 4.

As shown in Fig. 4, the distribution of δ_I obeyed the Rayleigh distribution $\frac{\delta_I}{\sigma_I^2}\exp\left(-\frac{\delta_I^2}{2\sigma_I^2}\right)$ [4–7], where σ_I is the standard deviation. The vertical broken line represents the average noise intensity $\overline{I_N} = \frac{1}{N}\sum_j \delta_I^j$, which takes non-zero value. At a signal, the distribution of $I(v)$ shifts to the right with a slight broadening and is represented by Nakagami–Rice distribution [4–7] (not shown). On the other hand, the distribution of δ_ϕ shows a more apparent depen-
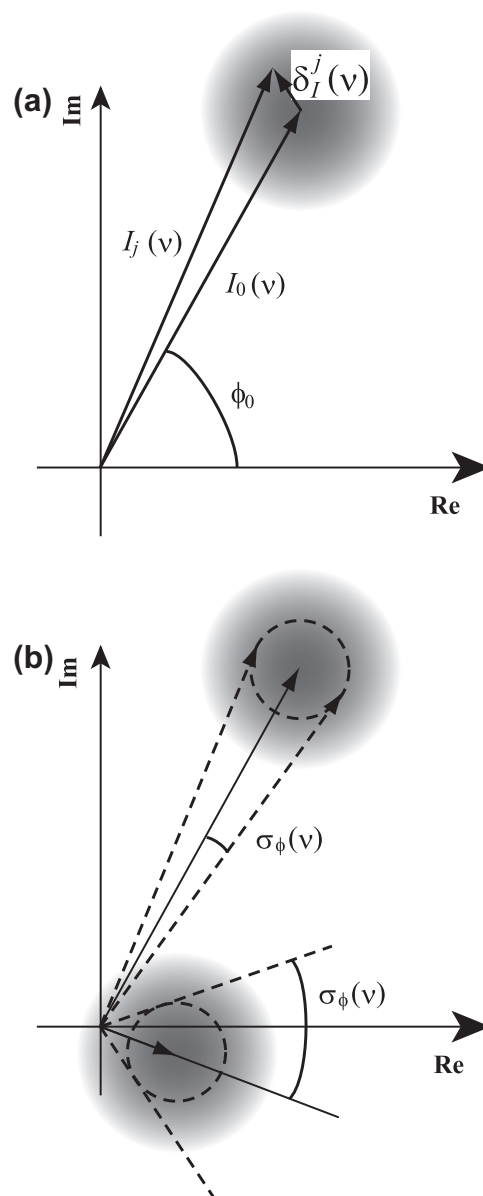

Fig. 1. (a) Schematic illustration of a complex data point $I_j(v) \cdot \exp(-i\phi_j(v))$, which is represented by a sum of a signal $I_0(v) \cdot \exp(-i\phi_0(v))$, and a noise $\delta_I^j(v) \cdot \exp(-i\delta_\phi^j(v))$. (b) Statistical distribution of the point represented by the standard deviation σ_ϕ of the phase characterizing the magnitude relation between I_0 and δ_I^j. The tone of the circle represents distribution of the noise.
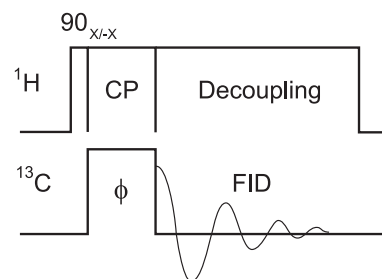
Fig. 2. Pulse sequence for ^{13}C high-resolution solid state NMR. The phase of the 90° pulse is altered by 180° for each spectral unit. When standard deviation spectrum is obtained, the phase ϕ of the ^{13}C CP phase is held in 0° and the FIDs with 90°x are subtracted from those with 90°−x in order to remove artifacts. When covariance spectrum is obtained, the ^{13}C CP phase ϕ was varied by, e.g. 1° in each of the two steps with the phase alternation of the 90° pulse.
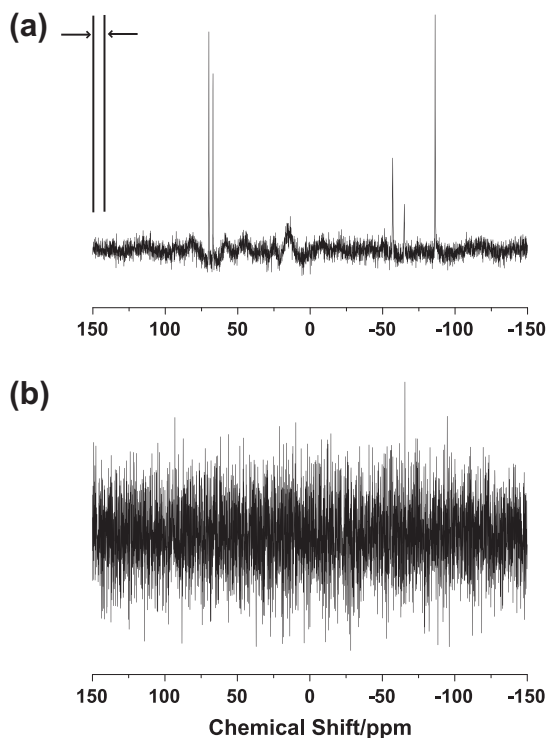
Fig. 3. ^{13}C CPMAS spectra of polycrystalline mixture of glycine and D,L-alanine obtained (a) by accumulating 1800 pairs of FIDs and (b) from a single pair of FID. The region indicated in (a) contains 100 data points, which are analyzed in the following discussion.
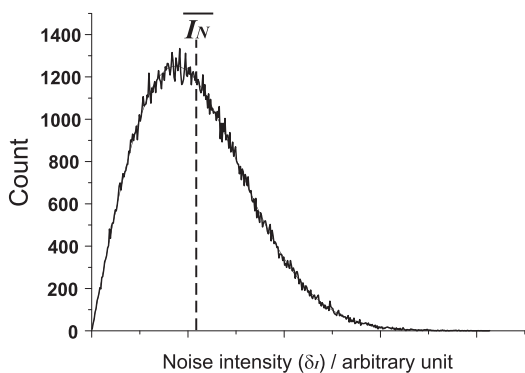
Fig. 4. A distribution of the intensity δ_I obtained from the 1800 × 100 data points in the noise region. $\overline{I_N}$ shown by dotted line represents the mean value of the intensity.

dence. Indeed, δ_ϕ is uniform over $-\pi \leqslant \delta_\phi \leqslant \pi$ for noise, while at signal δ_ϕ shows a broad single peak centered at ϕ_0 as will be shown in the next section.

Assuming that the intensity and the phase of the noise have no correlation, we evaluated statistically the signal intensity and phase individually. We compared the averaged absorption signal $\overline{I(v) \cdot \cos(\phi(v))}$ given by Eq. (6) (Fig. 3a) with the averaged signal intensity $\overline{I(v)}$ (Fig. 5a), the averaged phase $\overline{\cos(\phi(v))}$ (Fig. 5b), and their product $\overline{I(v)} \cdot \overline{\cos(\phi(v))}$ (Fig. 5c). The large peak of $\overline{I(v)}$ at the center is ascribed to the DC-offset. We found that separate accumulation of the intensity and the phase does not distort the spectrum significantly (compare Figs. 3a and 5c). It is interesting to note that $\overline{\cos(\phi(v))}$ was quite similar to the accumulated spectrum, but $\overline{I(v)}$ was not. This is because, as the SNR in a single spectrum was low in the present case (see Fig. 3b), the distribution of the intensity was dominantly determined by the noise and $\overline{I(v)}$ was
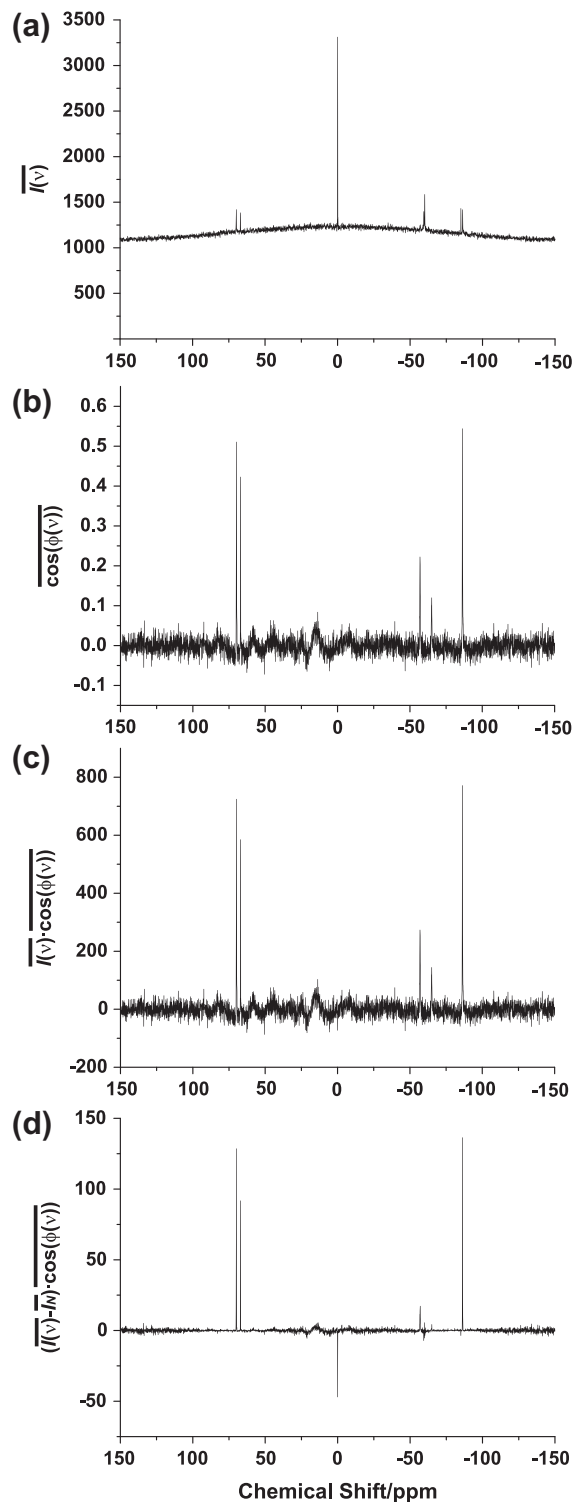
Fig. 5. Plots of (a) $\overline{I(v)}$, (b) $\overline{\cos(\phi(v))}$, (c) $\overline{I(v)} \cdot \overline{\cos(\phi(v))}$, and (d) $\overline{(I(v) - I_N)} \cdot \overline{\cos(\phi(v))}$.

similar at any v. In other words, when the SNR of each FID is low, $\overline{I(v)}$ is nearly constant and $\overline{I(v)} \cdot \overline{\cos(\phi(v))}$ is controlled by $\overline{\cos(\phi(v))}$. When the SNR of each FID is high, the contribution of $\overline{I(v)}$ to the lineshape $\overline{I(v)} \cdot \overline{\cos(\phi(v))}$ should be appreciable. For such a case, however, the tedious analysis of each FID for de-noising would not be necessary.

Here we examine in passing whether one can reduce noise by subtracting the average $\overline{I_N}$ of $I(v)$ from $\overline{I(v)}$ in the product. Fig. 5d shows $\overline{(I(v) - I_N)} \cdot \overline{\cos(\phi(v))}$, which has apparently better SNR.

However, the spectrum is significantly distorted and the small peaks at around −60 ppm are attenuated considerably.

2.3. Phase variance and phase standard deviation

In Section 2.1 we discussed the standard deviation σ_ϕ of the phase distribution for the two extreme cases of pure signal and pure noise. Here we derive a general formula of σ_ϕ for a given set of the signal intensity I_0 and the noise intensity σ_I. For this purpose we introduce a probability density function $H(\phi)$, which is the probability of finding the phase of v frequency point to be ϕ. Then, σ_ϕ can be represented as

$$\sigma_\phi^2 = \int_{-\pi}^{\pi} H(\phi(v))\phi(v)^2 d\phi(v) - \left(\int_{-\pi}^{\pi} H(\phi(v))\phi(v)d\phi(v)\right)^2. \quad (8)$$

In order to obtain $H(\phi)$, we firstly express noise in the Cartesian coordinate as $\delta_I^j(v) \cdot \exp(-i\delta_\phi^j(v)) = \delta_x^j(v) - i\delta_y^j(v)$, where $\delta_x^j(v)$ and $\delta_y^j(v)$ are the x(real)- and y(imaginary)-components of noise, respectively, and assume that both components follow the two-dimensional normal distribution given by

$$f_\delta(\delta_x(v), \delta_y(v)) = \frac{1}{2\pi\sigma_I^2} \exp\left(-\frac{\delta_x(v)^2}{2\sigma_I^2}\right) \exp\left(-\frac{\delta_y(v)^2}{2\sigma_I^2}\right). \quad (9)$$

Here, we use $\delta_x(v)$ and $\delta_y(v)$ as variables to express distribution of the x- and y-component of noise, respectively, and σ_I is the standard deviation of $f_\delta(\delta_x(v), \delta_y(v))$. Eq. (9) gives the probability density function whose origin is $(I_0(v), \phi_0(v))$. To calculate a probability density function $h(I(v), \phi(v))$ for $I_j(v)$ and $\phi_j(v)$, we shift the origin to $(0, 0)$ by substituting $\delta_x(v)$ and $\delta_y(v)$ as follows;

$$\delta_x(v) = I(v)\cos\phi(v) - I_0(v)\cos\phi_0(v) \quad (10)$$

and

$$\delta_y(v) = I(v)\sin\phi(v) - I_0(v)\sin\phi_0(v), \quad (11)$$

where $I(v)$ and $\phi(v)$ are the variables corresponding to $I_j(v)$ and $\phi_j(v)$, respectively. By putting Eqs. (10) and (11) into Eq. (9), we have

$$h(I(v), \phi(v)) = \frac{1}{2\pi\sigma_I^2} \exp\left(-\frac{(I(v)\cos\phi(v) - I_0(v)\cos\phi_0(v))^2}{2\sigma_I^2}\right)$$
$$\times \exp\left(-\frac{(I(v)\sin\phi(v) - I_0(v)\sin\phi_0(v))^2}{2\sigma_I^2}\right)$$
$$= \frac{1}{2\pi\sigma_I^2} \exp\left(-\frac{(I_0(v) - I(v)\cos\Delta\phi(v))^2 + (I(v)\sin\Delta\phi(v))^2}{2\sigma_I^2}\right), \quad (12)$$

where $\Delta\phi(v) = \phi(v) - \phi_0(v)$. Note here that the shift of the origin induces the apparent frequency dependence into the probability density function.

The probability density function for phase $H(\phi_j(v))$ is given by

$$H(\phi(v)) = \int_0^\infty I(v)h(I(v), \phi(v))dI(v)$$
$$= \frac{1}{2\pi\sigma_I^2} \exp\left(-\frac{I_0(v)^2}{2\sigma_I^2}\sin^2\Delta\phi_j(v)\right)$$
$$\times \int_0^\infty I(v) \cdot \exp\left(-\frac{(I(v) - I_0(v)\cos\Delta\phi_j(v))^2}{2\sigma_I^2}\right)dI(v)$$
$$= \frac{1}{2\pi} \exp\left(-\frac{R(v)^2}{2}\right) + \frac{R(v)\cos\Delta\phi_j(v)}{2\sqrt{2\pi}}$$
$$\times \exp\left(-\frac{R(v)^2}{2}\sin^2\Delta\phi_j(v)\right)$$
$$\cdot \left(1 + \mathrm{erf}\left(\frac{R(v)\cos\Delta\phi_j(v)}{\sqrt{2}}\right)\right), \quad (13)$$

where $R(v)$ is a ratio between the pure signal intensity $I_0(v)$ and the standard deviation of noise σ_I $R(v) = I_0(v)/\sigma_i$), which represents the SNR at v. The distribution depends solely on $R(v)$, and to appreciate its dependence, we calculated $H(\phi(v))$ for $R(v)$ = 0.1, 1.0, and 10 and plotted Fig. 6a–c, respectively. For small $R(v)$, $H(\phi(v))$ shows a broad distribution around $\phi_0(v)$ with the baseline determined by the first term in Eq. (13), and for $R(v) \gg 1$, $H(\phi(v))$ becomes a Gaussian-like distribution.

In Fig. 7, we plot $\sigma_\phi(v)$ calculated for several R values; for a pure noise $(R = 0)$, we have $\sigma_\phi(v) = \int_{-\pi}^{\pi} \frac{\phi^2}{2\pi}d\phi = \frac{\pi}{\sqrt{3}}$ (we define the value $\frac{\pi}{\sqrt{3}}$ as $\overline{\sigma_\phi^N}$ and use it below), which decays asymptotically to zero for $R = \infty$ (a pure signal).

Fig. 8 shows the distribution of $\phi(v)$ obtained from the experimental data used in the previous section. Fig. 8a and b shows the distribution of the phase at −86.2 ppm (a signal) and 150 ppm (noise). The distribution at the signal is similar to $H(\phi(v))$ calculated for R = 1.0 (Fig. 6b). This small R value is consistent with the poor SNR in each spectrum (Fig. 3b). The distribution of $\phi(v)$ for the noise area is flat as expected.
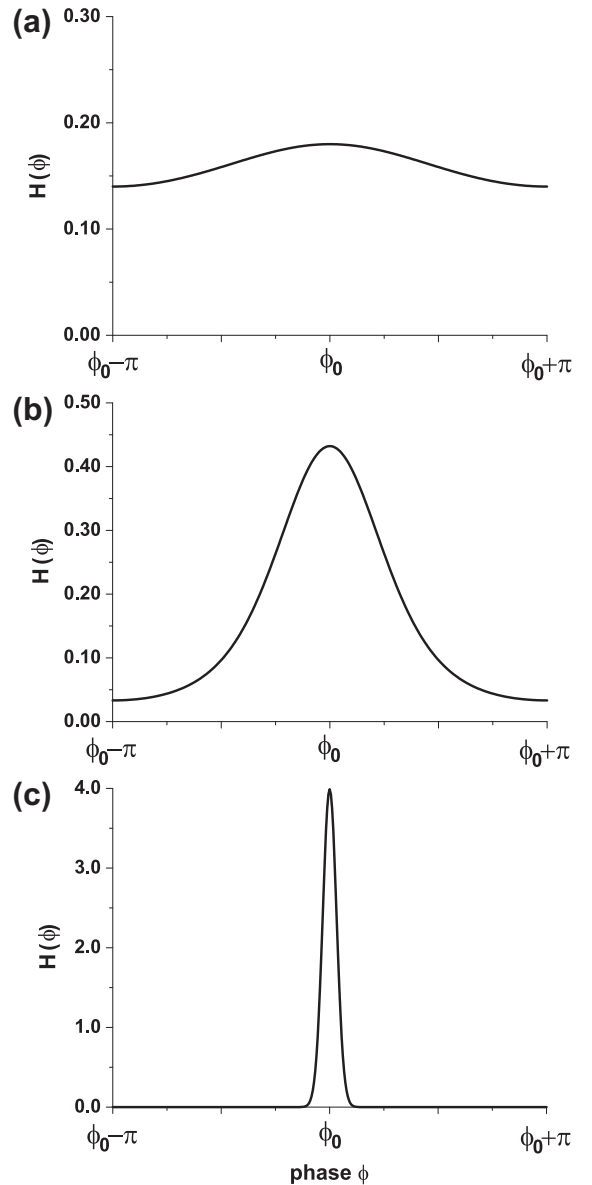

Fig. 6. The phase distribution function $H(\phi(v))$ for $R(v)$ = (a) 0.1, (b) 1.0, and (c) 10.

The standard deviation $\sigma_\phi(v)$ is calculated by using the distribution of $\phi(v)$ obtained for each v with $\phi_0(v)$ calculated from the accumulated spectrum at each point. To compare with the accumulated spectrum (Fig. 3a), we plot $1/\sigma_\phi(v)$ instead of $\sigma_\phi(v)$ in Fig. 9a. The $1/\sigma_\phi(v)$ plot resembles the accumulated spectrum. We then use $1/\sigma_\phi(v)$ as a weighting factor of the accumulated spectrum as $S(v) \cdot \left| \frac{1}{\sigma_\phi(v)} \right|$, which is shown in Fig. 9b. Apparently such weighting does not suppress noise. This is because the $1/\sigma_\phi(v)$ value at noise is large ca. 0.551 and is not significantly different from those at signals (0.595–0.878). Here we introduce a new weighting factor, referred to as a measure of certainty, which is plotted in Fig. 9c, and is defined as

$$q_{\sigma_\phi} = 1 - \frac{\sigma_\phi(v)}{\sigma_\phi^N} = 1 - \frac{\sqrt{3}\sigma_\phi(v)}{\pi}. \tag{14}$$

The measure of certainty q_{σ_ϕ} is zero for a noise and one for a pure signal. We then use the measure of certainty as a weighting factor of the accumulated spectrum as

$$S_{\text{PSDW}}(v) = S(v) \cdot |q_{\sigma_\phi}|, \tag{15}$$

which we shall refer to as a phase-standard-deviation weighted (PSDW) NMR spectrum, and the resulting spectrum is shown in Fig. 9d. Appreciable de-noising is achieved.

In the following, we compare the performance of the present analysis with the phase-covariance analysis [2]. However, before comparison, we shall briefly describe the phase-covariance analysis.

2.4. Phase covariance between the pulse and the signal

In the phase-covariance scheme [2], an array experiment is performed by incrementing the phase ψ_j of the rf pulse, and the correlation between the spectral and the rf phases is examined in terms of the covariance. A quadrature NMR signal at a j-th experiment may be given by introducing the rf phase factor of the j-th experiment ψ_j in Eq. (5) as

$$S_j(v) = \overline{I(v) \cdot \exp(-i(\phi(v) + \psi_j))}. \tag{16}$$

Similar to the conventional signal accumulation, the N FIDs are accumulated after taking the corresponding rf phase shift $\psi_j (j = 1 \cdots N)$ into account; the accumulated FID is Fourier transformed to produce the NMR spectrum $S(v)$. In addition to the conventional procedure, each FID signal is Fourier transformed separately to produce N NMR spectra $S_j(v_i)$. Practically, we determine the 0th and 1st order phase correlation values from the accu-
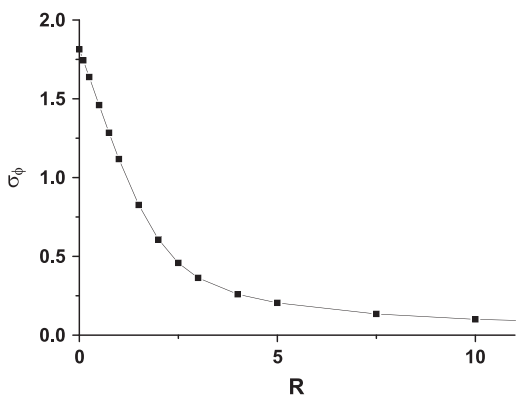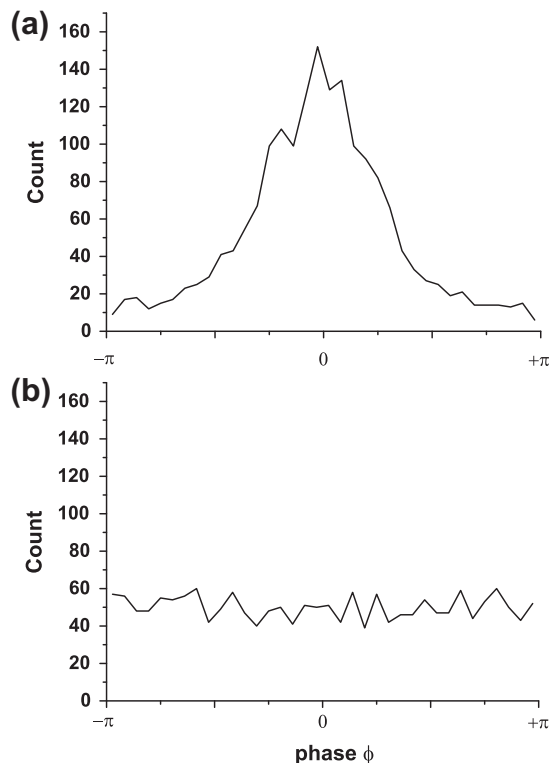

Fig. 8. The distribution of the phase $\phi(v)$ at (a) the peak, -86.2 ppm and (b) the noise area, around 150 ppm.

mulated spectrum $S(v)$. These phase values are used to correct each spectrum and the phase $p_j(v) = \phi(v) + \psi_j$ of each spectral point v in the j-th spectrum is deduced.

Fig. 10a shows two-dimensional distribution of $(p(v), \psi)$ for $\phi(v) = 0$ and $R(v) = 1.0$. Since $p_j(v)$ values obtained from experimental data include noise, observed $(p_j(v), \psi_j)$ pairs at the center of a signal $(\phi(v) = 0)$ would ideally distribute along the diagonal line. The cross-section at ψ_j is the phase probability density function $\eta(p_j(v), \psi_j)$ given by

$$\eta(p_j(v), \psi_j) = H(\phi(v) = p_j(v) - \psi_j)/2\pi. \tag{17}$$

The normalized covariance, which we call the correlation coefficient, is calculated for each point according to

$$\rho(v) = \frac{\frac{\sum_j p_j(v)\psi_j}{N} - \frac{\sum_j p_j(v)}{N} \cdot \frac{\sum_j \psi_j}{N}}{\sqrt{\left(\frac{\sum_j (p_j(v)^2)}{N} - \left(\frac{\sum_j p_j(v)}{N} \right)^2 \right) \cdot \left(\frac{\sum_j (\psi_j^2)}{N} - \left(\frac{\sum_j \psi_j}{N} \right)^2 \right)}}. \tag{18}$$

Practically, however, one apparent problem may be noted, that is, the $(p_j(v), \psi_j)$ data at the corners $(-\pi, +\pi)$ and $(+\pi, -\pi)$ in Fig. 10a does reduce the $\sum_j p_j(v)\psi_j$ term and thereby correlation coefficient. The $(p_j(v), \psi_j)$ data at the corners arise because of the periodicity of angles. Instead of developing a theory for evaluating a correlation factor in such a case, we transform $(p_j(v), \psi_j)$ to $(\Phi_j(v), \Psi_j)$ as
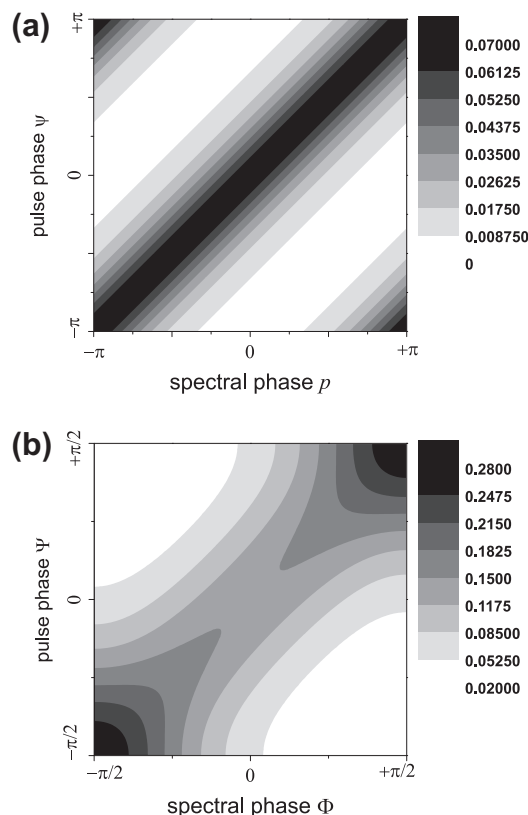
$$\Phi_j(v) = |p_j(v)| - \frac{\pi}{2} \tag{19}$$

and

$$\Psi_j = |\psi_j| - \frac{\pi}{2}. \tag{20}$$

Fig. 10b illustrates the result of the transformation. The phase probability density function in this region $g(\Phi_j(v), \Psi_j)$ is written as

Fig. 7. R dependence of the standard deviation $\sigma_\phi(v)$ of the phase.

(a)

(b)

(c)

(d)
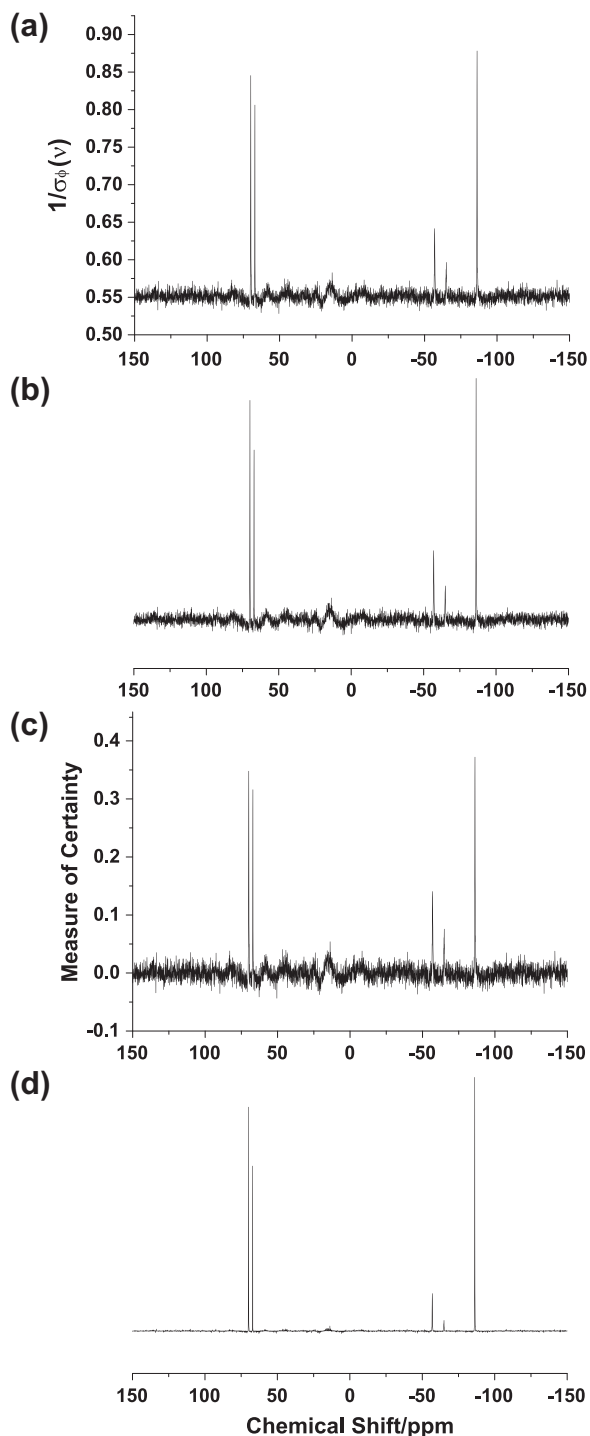
Fig. 9. Comparison of (a) a plot of the inverse of the standard deviation $1/\sigma_\phi(v)$, (b) the $1/\sigma_\phi(v)$ weighted spectrum $S(v) \cdot |\frac{1}{\sigma_\phi(v)}|$, (c) a plot of the measure of certainty of the standard deviation q_{σ_ϕ}, and (d) the PSDW NMR spectrum.

$$
\begin{aligned}
g(\Phi_j(v), \Psi_j) = {} & \eta\left(\Phi_j(v) + \frac{\pi}{2}, \Psi_j + \frac{\pi}{2}\right) \\
& + \eta\left(-\Phi_j(v) - \frac{\pi}{2}, \Psi_j + \frac{\pi}{2}\right) \\
& + \eta\left(\Phi_j(v) + \frac{\pi}{2}, -\Psi_j - \frac{\pi}{2}\right) \\
& + \eta\left(-\Phi_j(v) - \frac{\pi}{2}, -\Psi_j - \frac{\pi}{2}\right).
\end{aligned} \quad (21)
$$

(a)

(b)

Fig. 10. (a) Distribution fraction of the pulse phase ψ and signal phase $p(v)$ and (b) that of ψ and $\phi(v)$ calculated according to Eqs. (19) and (20) in the case of $\phi_0(v) = 0$ or $\phi_0(v) = 0$, and $R(v) = 1.0$.

The two-dimensional phase field $(\Phi(v), \Psi)$ corresponds to a $(p(v), \psi)$ square folded along the $p(v) = 0$ and $\psi = 0$ lines. We refer to this field as the folded field. The correlation coefficient $\rho_f(v)$ in the folded field is written as

$$
\rho_f(v) = \frac{\frac{\sum_j \Phi_j(v)\Psi_j}{N} - \frac{\sum_j \Phi_j(v)}{N} \cdot \frac{\sum_j \Psi_j}{N}}{\sqrt{\left(\frac{\sum_j (\Phi_j(v)^2)}{N} - \left(\frac{\sum_j \Phi_j(v)}{N}\right)^2\right) \cdot \left(\frac{\sum_j (\Psi_j^2)}{N} - \left(\frac{\sum_j \Psi_j}{N}\right)^2\right)}}. \quad (22)
$$

In the phase-covariance analysis, we use the absolute value of $\rho_f(v)$ as a weighting factor at v as

$$
S_{PCW}(v) = |\rho_f(v)|S(v), \quad (23)
$$

which we call a phase-covariance weighted (PCW) NMR spectrum.

2.5. Comparison of the phase-standard-deviation analysis and the phase-covariance analysis

To examine the de-noising effect for both analyses, we carried out ^{13}C CPMAS experiments by intentionally adding an incoherent noise to the receiver circuit. The receiver cable was coupled with an additional signal from a frequency synthesizer which was not synchronized with the NMR spectrometer. Further, to examine the effects of these de-noising analyses to a small signal, we used a shorter delay time of 10 s. As a consequence, the relative intensity of glycine peaks was much smaller than that in previous sections. One of the paired-scan spectra is given in Fig. 11a, in which the strong peak at 33.2 ppm is the introduced noise. The ^{13}C CP phase ϕ (Fig. 2) was varied by 1° step in each of the paired FIDs with the 90° pulse phase alternation. For the PSDW analysis,
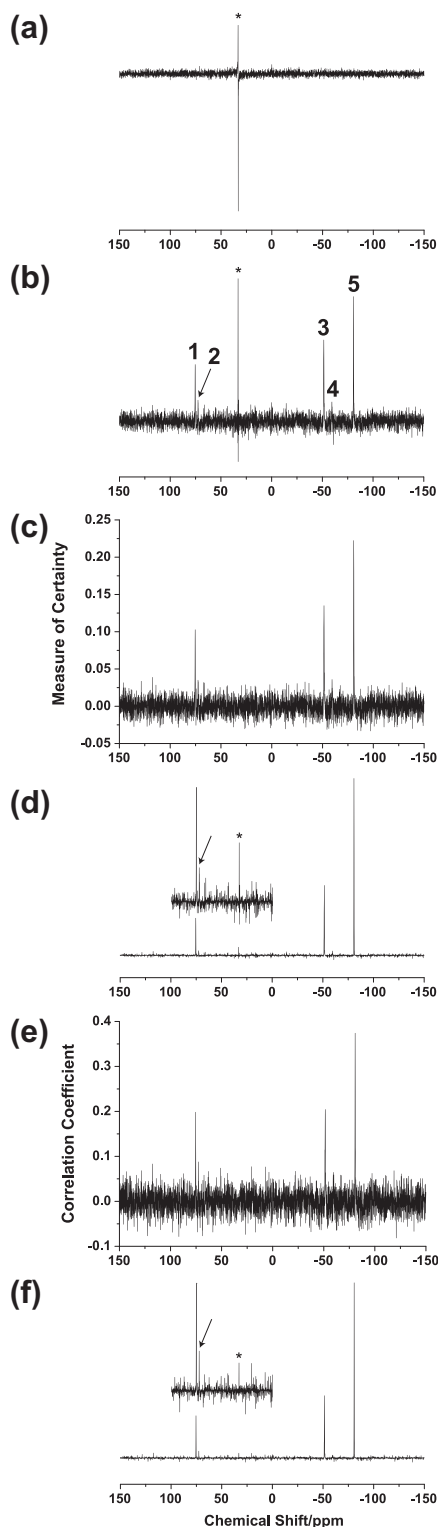
Fig. 11. ^{13}C CPMAS spectra of polycrystalline mixture of glycine and L-alanine measured by adding an additional incoherent frequency signal at 33.2 ppm. (a) The spectrum obtained from a single FID. (b) The accumulated spectrum. (c) The measure of certainty spectrum. (d) The PSDW NMR spectrum. (e) The correlation coefficient spectrum. (f) The PCW NMR spectrum. The asterisks and arrows indicate the noise at 33.2 ppm and the C=O signal of glycine at 72.6 ppm, respectively.

these 1° phase shifts were compensated by phase correction, while for the PCW analysis, they were left to appreciate correlation. It is worth nothing here that PSDW and PCW, as well as signal averag-
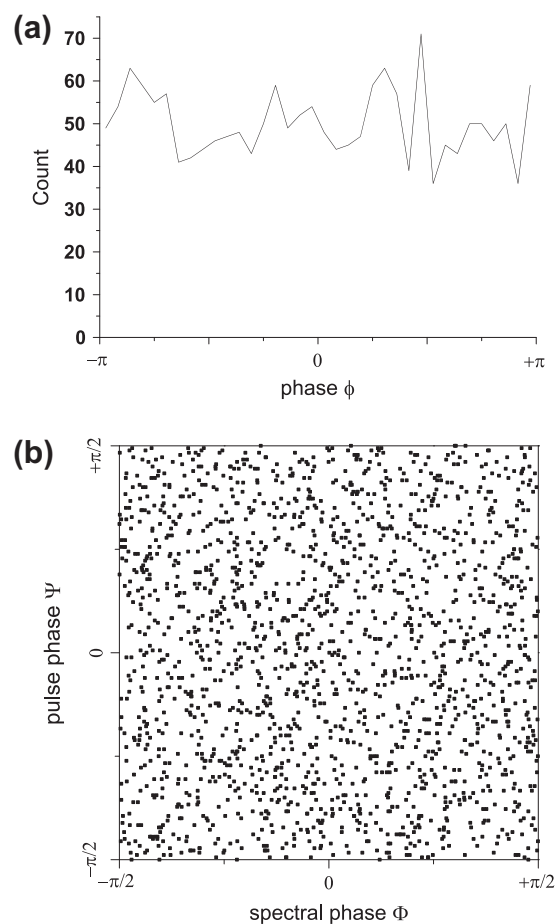
Fig. 12. (a) A distribution plot of $\phi(v)$ and (b) a distribution pattern of $(\Phi(v), \Psi)$ at 33.2 ppm.

Table 1
Peak area intensities in the $S_W(v)$, PSDW, and PCW spectra shown in Fig. 11. The peak numbers correspond to those indicated in Fig. 11b, and the area intensities are normalized by that of the incoherent noise (*).

Peak	1	2	3	4	5
$S_W(v)$	2.030	0.417	4.302	0.978	4.484
PSDW	19.90	2.050	50.19	2.861	84.85
PCW	22.21	2.070	51.10	3.443	88.28
PSDW/$S_W(v)$	9.803	4.916	11.67	2.925	18.92
PCW/$S_W(v)$	10.94	4.964	11.88	3.520	19.69

ing, use a common data array, so that all of these post processing can be performed concurrently.

Fig. 11b shows the accumulated spectrum. Even with 1800 pairs of FIDs, the noise introduced at 33.2 ppm still exists. Fig. 11c and e are the measure of certainty and correlation coefficient, i.e., the weighting factors of the PSDW spectrum and PCW spectrum, respectively, and Fig. 11d and f are the PSDW spectrum and the PCW spectrum. In both methods, the incoherent noise was effectively eliminated. As shown in the magnified plots of the spectra, the PCW resulted in better suppression of the incoherent noise than PSDW. This may be explained as follows. Fig. 12a and b shows the distribution of $\phi(v)$ and the distribution pattern of $(\Phi(v), \Psi)$ for the huge noise at 33.2 ppm. In PSDW, the number of pairs of scans ($N = 1800$) is not large enough to give a flat distribution of $\phi(v)$. On the other hand, the distribution pattern of $(\phi(v), \psi)$ is comparatively more uniform, leading to a small correlation coefficient in
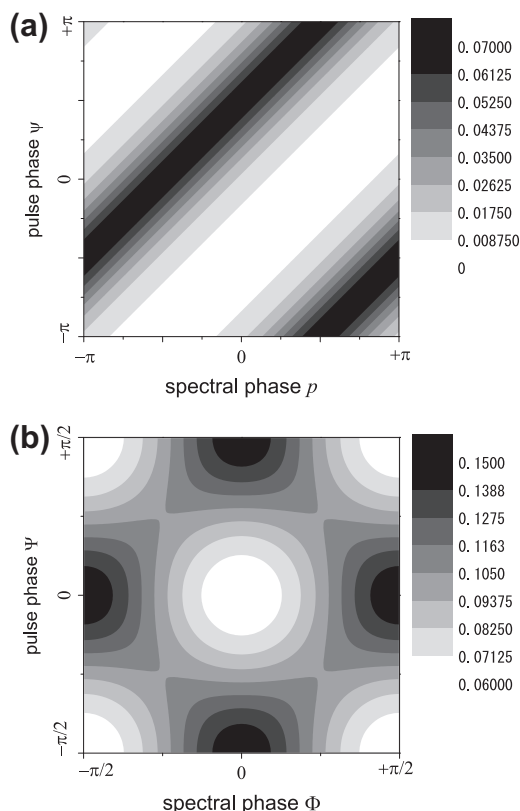
Fig. 13. (a) Distribution fraction of the pulse phase ψ and signal phase $p(v)$ and (b) that of Ψ and $\Phi(v)$ calculated according to Eqs. (19) and (20) in the case of $\phi_0(v) = \pi/2$ or $\Phi_0(v) = \pi/2$, and $R(v) = 1.0$.

the PCW approach. It follows that the phase-covariance analysis is more robust for relatively small number of data arrays.

For the small C=O signal of glycine at 72.6 ppm (peak 2), designated by an arrow in Fig. 11, the peak was apparently reduced by both methods. Nevertheless, its relative intensity to that of the incoherent noise was still larger than in the case of the conventional accumulation. Table 1 shows integrated peak normalized by that of the incoherent noise. Since the weighted spectra correspond to the power spectrum, we consider $S(v)|S(v)| \equiv S_w(v)$ for the accumulated spectrum. In PSDW and PCW spectra, the large peak 1, 3, and 5 are noticeably enlarged, the largest peak 5 become about 20 times, and even the small peak 2 and 4, which are smaller than the peak of incoherent noise peak in $S_w(v)$ spectrum, are enough larger than incoherent noise peak, the peak 4 become about three times. PSDW and PCW distort relative intensities, but they are useful to judge signals and noises. More quantitative statistical analysis on SNR of PSDW and PCW is underway and will be published elsewhere.

Lastly, we point out two problems, which arise when analyzing the intensity and the phase separately. The first one of them relates to the use of the absolute intensity $I(v)$ instead of the absorption-corresponding lineshape. Since the former has a wider linewidth than the latter, the measure of certainty for noises on the envelope

of a strong signal tends to be overestimated. For example, at the frequency where the absorption intensity decays to 10%, the absolute intensity decays only to 22%, resulting in a non-flat distribution for $H(\phi(v))$. The second one relates to the evaluation of the effect of a non-zero $\phi_0(v)$ value when calculating the standard deviation. This problem of the non-zero ϕ_0 is apparent for the phase-covariance analysis with the folding as shown in Fig. 13, illustrating for the case of $\phi_0 = \pi/2$ the distribution for the full (Fig. 13a) and the folded (Fig. 13b) fields, respectively. Since the intensity for an absorption lineshape in the region of $\phi_0 = \pm\pi/2$ is less prominent, the apparent correlation coefficient calculated from the former case would leave noises in this region. The folding makes the correlation coefficient at $\phi_0 = \pi/2$ small (ideally zero) and gives reasonable coefficient values at the whole envelope region.

3. Conclusions

In this work, we examined two statistical approaches, namely, phase standard deviation weighting (PSDW) and phase covariance weighting (PCW) analysis for de-noising, and as a result the usefulness of their methods is verified. These methods are applicable to many situations, regardless of the state of sample: solid or solution, the dimension of the spectrum: 1D, 2D or more, and the pulse sequence. The present methods can be used with other de-noising techniques [8,9]. PSDW and PCW distort relative intensities of signals favoring larger peaks, and cannot be applied to lineshape analysis or quantitative analysis. Nevertheless, they are useful techniques to judge signals and noises, as demonstrated in the elimination of the incoherent spurious signal. When broadcasting radio waves contaminate NMR spectra, the present approach can be a solution to get rid of them. Details of further examination using simulations to show whether the limits of visible signal in PSDW and PCW are weaker than that in classic accumulation or not, data processing, and discussion of applicability to more sophisticated pulse sequences will be presented elsewhere.

References

[1] N. Ivchenko, C.E. Hughes, M.H. Levitt, J. Magn. Reson. 160 (2003) 52–58.
[2] J. Fukazawa, K. Takegoshi, Phys. Chem. Chem. Phys. 12 (2010) 11225–11227.
[3] D.I. Hoult, R.E. Richards, Proc. R. Soc. London, Ser. A 344 (1975) 311–340.
[4] J. Sijbers, A.J. den Dekker, J. Van Audekerke, M. Verhoye, D. Van Dyck, Magn. Reson. Imag. 16 (1998) 87–90.
[5] J. Sijbers, A.J. den Dekker, E. Raman, D. Van Dyck, Int. J. Imag. Syst. Technol. 10 (1999) 109–114.
[6] M. Nakagami, The m-Distribution, a general formula of intensity of rapid fading, in: W. C. Hoffman (Ed.), Statistical Methods in Radio wave Propagation: Proceedings of A symposium held at the University of California, Los Angeles, June 18–20, 1958, Pergamon Press, London, 1960, pp. 3–36.
[7] S.O. Rice, Bell Syst. Tech. J. 24 (1945) 46–156.
[8] J.B. Weaver, D. Healy Jr., J. Magn. Reson., Ser. A 113 (1995) 1–10.
[9] G. Sarty, E. Kendall, J. Magn. Reson., Ser. B 111 (1996) 50–60.