The Journal of Chemical Physics

Optimized "detectors" for dynamics analysis in solid-state NMR

Cite as: J. Chem. Phys. **148**, 045104 (2018); https://doi.org/10.1063/1.5013316
Submitted: 13 November 2017 . Accepted: 09 January 2018 . Published Online: 29 January 2018

Albert A. Smith, Matthias Ernst, and Beat H. Meier

AIP Publishing

Optimized "detectors" for dynamics analysis in solid-state NMR

Albert A. Smith, Matthias Ernst,[a] and Beat H. Meier[b]
ETH Zurich, Physical Chemistry, Vladimir-Prelog-Weg 2, 8093 Zurich, Switzerland

(Received 13 November 2017; accepted 9 January 2018; published online 29 January 2018)

Relaxation in nuclear magnetic resonance (NMR) results from stochastic motions that modulate anisotropic NMR interactions. Therefore, measurement of relaxation-rate constants can be used to characterize molecular-dynamic processes. The motion is often characterized by Markov processes using an auto-correlation function, which is assumed to be a sum of multiple decaying exponentials. We have recently shown that such a model can lead to severe misrepresentation of the real motion, when the real correlation function is more complex than the model. Furthermore, multiple distributions of motion may yield the same set of dynamics data. Therefore, we introduce optimized dynamics "detectors" to characterize motions which are linear combinations of relaxation-rate constants. A detector estimates the average or total amplitude of motion for a range of motional correlation times. The information obtained through the detectors is less specific than information obtained using an explicit model, but this is necessary because the information contained in the relaxation data is ambiguous, if one does not know the correct motional model. On the other hand, if one has a molecular dynamics trajectory, one may calculate the corresponding detector responses, allowing direct comparison to experimental NMR dynamics analysis. We describe how to construct a set of optimized detectors for a given set of relaxation measurements. We then investigate the properties of detectors for a number of different data sets, thus gaining an insight into the actual information content of the NMR data. Finally, we show an example analysis of ubiquitin dynamics data using detectors, using the DIFRATE software. *Published by AIP Publishing.* https://doi.org/10.1063/1.5013316

I. INTRODUCTION

Nuclear Magnetic Resonance (NMR) is a powerful method for characterizing both molecular structure and dynamics. The latter application is often used to determine local dynamics of molecules, especially proteins, via the measurement of multiple relaxation-rate constants. Those relaxation-rate constants depend on the internal motion as well as rigid-body motion of the protein[1] and in the case of solution-state NMR, the overall tumbling of the protein.[2,3] The analysis of relaxation-rate constants in solid-state NMR typically relies on modeling the correlation function using a sum of multiple exponential functions each characterized by a correlation time and an amplitude. This was inspired by the success of the Lipari-Szabo model-free approach[4] or extensions of that approach[5–7] in solution-state NMR, achieved by setting the correlation time of the overall tumbling in solids to infinity. While the validity of the Lipari-Szabo model-free approach in solution-state NMR was justified in detail (similar correlation functions are derived by Halle and Wennerström,[8] with the validity of this and the model-free approach discussed in detail[9]), studies in solid-state NMR have investigated primarily the effects that the model selection[10–12] and the data included[13] have on the analysis of the experimental data. Although there are indications that modeling the correlation function with a sum of exponential functions is not always reliable, these studies do not directly consider modeling behavior in the case that the real motion is too complex to be fully characterized by the experimental data. In a theoretical study, we have recently shown that dynamics analysis can deviate significantly from the real characteristics of the motion if the real internal correlation function is more complex (i.e., if it has more correlation times) than the model used.[14] For example, if one uses a model correlation function with two exponential functions and the real motion is described by a tri-exponential correlation function, the resulting amplitudes and correlation times from the analysis may deviate significantly from the real motion. Furthermore, we have shown that many different distributions of motion can result in identical sets of relaxation data so that dynamics data acquired with NMR are inherently ambiguous in its description of dynamics.

These problems lead us to consider an alternative approach to analyze relaxation data, which takes into consideration only the information actually contained in the experimental data. We introduce the concept of a dynamics *detector*, which reports the average or total amplitude of motion within a range of correlation times that is defined by the sensitivity of that detector. A dynamics analysis would then report *responses* of multiple detectors, which contain information about the amplitude of the dynamics in different ranges of correlation times. Detectors are particularly powerful when comparing dynamics information obtained from NMR relaxation experiments to other methods, both because the sensitivity of the detectors clearly communicates the available information

[a]maer@ethz.ch. Telephone: +41 44 6324366.
[b]beme@ethz.ch. Telephone: +41 44 6324401.

from the NMR experiments and the detector responses are well-defined quantities; that is, they may be calculated if one has a distribution of the motion, which may be obtained from a trajectory of molecular motion (e.g., molecular-dynamics simulation). Thus, a detector analysis provides the natural link between experimental NMR data and computational methods. Note that the analysis method presented here was in part inspired by the theory of color vision.[15,16]

In the following, we describe how one designs dynamics detectors for a given set of solid-state NMR relaxation data and investigate the properties of those detectors for several types of data sets, concluding with a dynamics analysis of the model protein ubiquitin. Similar procedures can also be implemented for solution-state NMR. Since there are, however, important differences in the way the detectors are constructed, we will discuss the solution-state NMR approach in a separate publication.

II. RELAXATION FORMALISM

In this study, we will treat relaxation that is the result of the stochastic re-orientational motion of molecules that modulate anisotropic NMR interactions. This motion can be described by a correlation function, which we assume to be a sum of decaying exponentials. For an arbitrary number of exponentials, the correlation function can be represented by

$$\begin{aligned} C(t) = \frac{1}{5}\Big[&(1-S_1^2)\exp(-t/\tau_1) + S_1^2(1-S_2^2)\exp(-t/\tau_2) \\ &+ S_1^2S_2^2(1-S_3^2)\exp(-t/\tau_3) + \cdots + S_1^2S_2^2S_3^2\ldots\Big], \\ &\tau_1 < \tau_2 < \tau_3 < \cdots, \end{aligned} \tag{1}$$

where the factor 1/5 comes from the term $1/(2\ell+1)$ in the correlation function, where ℓ is the rank of the interaction tensor (in our case, rank-2).[17,18] Note that in this form, we assume that the motions are independent and the total correlation function is a product of the correlation functions for each motion. Then, if τ_n are not well-separated, they should be replaced by effective correlation times.[17] In a more compact form, Eq. (1) can be written as

$$C(t) = \frac{1}{5}\left[S^2 + (1-S^2)\sum_{i=1} A_i \exp(-t/\tau_i)\right], \tag{2}$$

where

$$\begin{aligned} &(1-S^2)A_1 = (1-S_1^2), \\ &(1-S^2)A_2 = S_1^2(1-S_2^2), \\ &(1-S^2)A_3 = S_1^2S_2^2(1-S_3^2), \ldots, \\ &\sum_{i=1} A_i = 1, \\ &\prod_{i=1} S_i^2 = S^2. \end{aligned} \tag{3}$$

Equation (2) is general, even in the case that τ_i in Eq. (1) are not well-separated, although the values of τ_i will be replaced by effective correlation times in both equations. We note that Eq. (2) allows for an infinite number of exponentials so that we can also express the correlation function as a distribution of correlation times. If we take $(1–S^2)\theta(z)$ to be our distribution of motion, then

$$C(t) = \frac{1}{5}\left[S^2 + (1-S^2)\int_{-\infty}^{\infty}\theta(z)\exp(-t/(10^z\cdot 1\text{ s}))dz\right], \tag{4}$$

where we describe the distribution on a logarithmic scale,[19] with

$$\begin{aligned} z &= \log_{10}(\tau_c/1\text{ s}), \\ \int_{-\infty}^{\infty}\theta(z)dz &= 1. \end{aligned} \tag{5}$$

The spectral density function, $J(\omega)$, can be obtained by the Fourier transformation of the correlation function, yielding

$$J(\omega) = \frac{2}{5}(1-S^2)\int_{-\infty}^{\infty}\theta(z)\frac{10^z\times 1\text{ s}}{1+(\omega\times 10^z\times 1\text{ s})^2}dz. \tag{6}$$

The various relaxation-rate constants can be calculated from the spectral-density function as linear combinations sampled at different frequencies.[2,3] We give the analytical expressions for R_1 and $R_{1\rho}$ as examples—see Ref. 20 for a more comprehensive discussion. The longitudinal relaxation, R_1, is given by

$$\begin{aligned} R_1 = &\left(\frac{\delta_{IS}}{4}\right)^2 (J(\omega_I-\omega_S) + 3J(\omega_I) + 6J(\omega_I+\omega_S)) \\ &+ \frac{3}{4}(\omega_I\sigma_{zz})^2 J(\omega_I), \end{aligned} \tag{7}$$

where relaxation is measured on spin I and is induced by a dipolar coupling to spin S (characterized by the anisotropy δ_{IS}) and the chemical-shift anisotropy (CSA) of spin I (characterized by $\omega_I\sigma_{zz}$). Here, ω_I and ω_S are the nuclear Larmor frequencies of I and S spins, respectively, in angular units.

In solid-state NMR, on-resonance rotating-frame relaxation for a heteronuclear spin pair under magic-angle spinning (MAS) is given by

$$\begin{aligned} R_{1\rho} = &\frac{1}{2}R_1 + \left(\frac{\delta^{IS}}{4}\right)^2\left(3J(\omega_S) + \frac{1}{3}J(\omega_1-2\omega_r) + \frac{2}{3}J(\omega_1-\omega_r)\right. \\ &\left. + \frac{2}{3}J(\omega_1+\omega_r) + \frac{1}{3}J(\omega_1+2\omega_r)\right) \\ &+ \frac{1}{6}(\omega_I\sigma_{zz})^2\left(\frac{1}{2}J(\omega_1-2\omega_r) + J(\omega_1-\omega_r) + J(\omega_1+\omega_r)\right. \\ &\left. + \frac{1}{2}J(\omega_1+2\omega_r)\right), \end{aligned} \tag{8}$$

where ω_r and ω_1 are the MAS frequency and the spin-lock field-strength, in angular-frequency units[21,22] (for a homonuclear spin-pair, see Ref. 23). Although other relaxation experiments are possible, only R_1 and $R_{1\rho}$ data will be used as examples in this study.

For convenience, we define a functional form of the relaxation-rate constants, denoted $R_\zeta(z)$, where ζ indicates the particular experiment and conditions (for example, R_1 at an external field of 500 MHz would be denoted $R_{1,500}(z)$). Then, $R_\zeta(z)$ is the relaxation-rate constant calculated for a mono-exponential correlation function, having a correlation time

$\tau_c = 10^z$ s and an order parameter $1 - S^2 = 1$. Using this functional form of the relaxation-rate constants, it is possible to calculate the rate constant for any arbitrary distribution, given by $(1 - S^2)\theta(z)$, as

$$R_\zeta^{(\theta,S)} = (1 - S^2) \int_{-\infty}^{\infty} \theta(z) R_\zeta(z) dz. \tag{9}$$

We will refer to $R_\zeta(z)$ as the *sensitivity* of the rate constant, while $R_\zeta^{(\theta,S)}$ is the relaxation-rate constant resulting from a distribution of motion characterized by $\theta(z)$ and $1 - S^2$.

III. DETECTORS IN SOLID-STATE NMR RELAXATION

Figure 1(a) plots sensitivities of the ^{15}N R_1 relaxation-rate constant at ^{1}H Larmor frequencies of 400 and 850 MHz. Each rate constant provides some information on the internal motion, but the sensitivities of both relaxation-rate constants ($R_{1,400}(z)$, $R_{1,850}(z)$) are rather broad and overlapping so that one does not get very precise information about which range of correlation times is being detected. Furthermore, the sensitivities of the two rate constants cover a similar range of correlation times so that without further processing, the additional information provided by the second rate constant is convoluted with the information already obtained from the first rate constant.

Although the rate constants have a similar range of sensitivity, they do not have the same functional form so that it is possible to separate information obtained from different time scales. To achieve this, we define a *detector response*, which is obtained from a linear combination of rate constants. In this example, the detector responses are given by

$$\begin{aligned} \rho_1^{(\theta,S)} &= aR_{1,400}^{(\theta,S)} + bR_{1,850}^{(\theta,S)}, \\ \rho_2^{(\theta,S)} &= cR_{1,400}^{(\theta,S)} + dR_{1,850}^{(\theta,S)}. \end{aligned} \tag{10}$$

We may similarly define a *detector sensitivity*, which is obtained using the same linear combination, but now for the sensitivities of the two rate constants,

$$\begin{aligned} \rho_1(z) &= aR_{1,400}(z) + bR_{1,850}(z), \\ \rho_2(z) &= cR_{1,400}(z) + dR_{1,850}(z). \end{aligned} \tag{11}$$

Note that we define the factors *a-d* to have units of seconds so that the resulting detector responses and detector sensitivities are unitless. Because of the linearity, we can also calculate the detector responses according to the following equation,

$$\rho_n{}^{(\theta,S)} = (1 - S^2) \int_{-\infty}^{\infty} \theta(z) \rho_n(z) dz, \tag{12}$$

so that the detector responses measure the overlap of the distribution of motion, $(1 - S^2)\theta(z)$, with the sensitivity of each detector, $\rho_n(z)$. The supplementary material has a glossary of all the different quantities defined in this section.

One then optimizes a, b, c, and d so that the sensitivities of the two detectors, $\rho_1(z)$ and $\rho_2(z)$, are as separated as possible, without having any negative values for each $\rho_n(z)$. The latter is an important requirement because we want to avoid the situation where motion at a correlation time with negative sensitivity and motion at a correlation time with positive sensitivity could cancel out. We achieve this by requiring the ratio of −b/a to be given by $\max(R_{1,400}(z)/R_{1,850}(z))$ and −d/c to be given by $\min(R_{1,400}(z)/R_{1,850}(z))$ and scaling such that $\rho_n(z)$ have maxima of 1. Our optimized detectors, $\rho_n(z)$, are shown in Fig. 1(b) with $\rho_1(z)$ covering the fast motions, $\rho_2(z)$ covering the slower motions, and some overlap in the center around $\tau_c = 10^{-8.5}$ s. Then, a protein motion with a large $\rho_1^{(\theta,S)}$ and a small $\rho_2^{(\theta,S)}$ value would be indicative of a distribution that has most motion with correlation times shorter than $\sim 10^{-9}$ s. The converse would indicate that most motion has correlation times longer than $\sim 10^{-8}$ s. Similar values of $\rho_1^{(\theta,S)}$ and $\rho_2^{(\theta,S)}$ indicate either motion in between these correlation times or similar amounts of motion above and below the two correlation times.

Without additional information or making assumptions, the responses of the two detectors summarize the actual information content of the data without bias, whereas an explicit model may be biased towards particular correlation times. In the following discussion, we will present a general strategy for constructing detectors in the case of many relaxation measurements and investigate the properties of those detectors. Note that linear combinations of rate constants have already been used to characterize motion in the spectral-density mapping approach to dynamics analysis.[24–27] However, the method presented here is a more general approach, with the intended goal of characterizing the distribution of motional time scales, $(1 - S^2)\theta(z)$, as opposed to the spectral density, $J(\omega)$. The relationship of the two methods is discussed below (see Sec. III D).

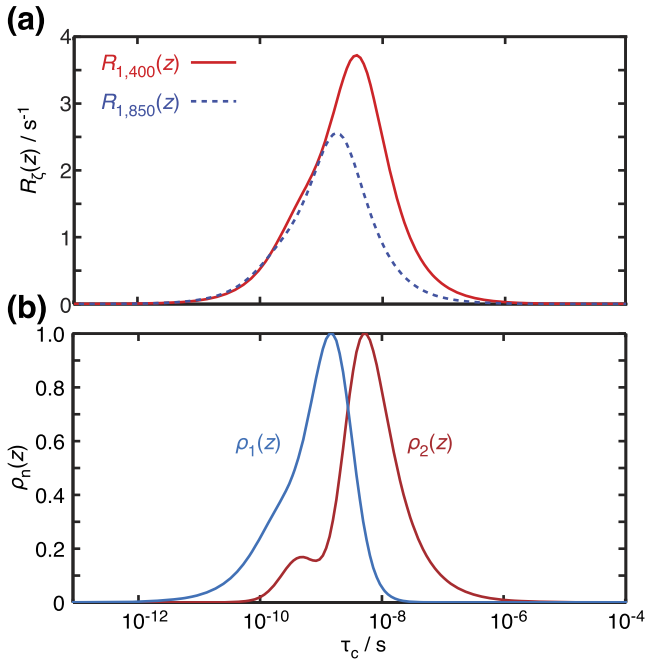

FIG. 1. ^{15}N R_1 relaxation rate constants at 400 MHz and 850 MHz. The R_1 sensitivities plotted in (a) are calculated for motions having a single correlation time, plotted on the x-axis, and an amplitude of $1 - S^2 = 1$. In (b), the sensitivity of two detectors, $\rho_n(z)$, are plotted that have been constructed by linear combinations of the R_1 rate constants. The coefficients used to obtain $\rho_n(z)$ are $a = -0.2017$ s, $b = 0.6189$ s, $c = 0.4607$ s, and $d = -0.3917$ s.

A. Two relaxation rate constants

We further develop the idea of detectors for the analysis of NMR relaxation data, continuing with the example of two ^{15}N R_1 relaxation-rate constants since for such an example, a graphical representation of the space of allowed relaxation data is a simple two-dimensional (2D) plot. To understand the ambiguity in relaxation-rate measurements and how careful detector construction can maximize information about the time scale of motions without making assumptions about the model of motion, we introduce the *allowed* region of relaxation-rate constants.

The allowed region of rate constants for ^{15}N R_1 at 400 and 850 MHz is shown in Fig. 2 highlighted in light blue. This allowed region is defined as any pair of rate constants that can be produced by some arbitrary distribution of motion, $(1-S^2)\theta(z)$. Because the two rate constants are both dependent on the same distribution of correlation times, $\theta(z)$, and total amplitude, $1-S^2$, only certain combinations of the rate constants are possible so that only part of the space is covered. If two distributions of motions (given by $(1-S^2)\theta(z)$) are experimentally distinguishable, they fall on different points in the allowed region, whereas if they are not, they fall on the same point. The distance between two points indicates how easily they can be distinguished which might depend on the available precision of the experimental measurements. It is easy to see that two different distributions of motions may yield the same point in the space, i.e., lead to the same pair of relaxation-rate constants. For example, colored traces in the allowed region plot out pairs of $R_{1,400}^{(\theta,S)}$ and $R_{1,850}^{(\theta,S)}$ values for motions with only a single correlation time ($\theta(z)$ is a δ-function), for several values of $1-S^2$. Every point in the allowed region can be the result of such a motion having only a single correlation time.

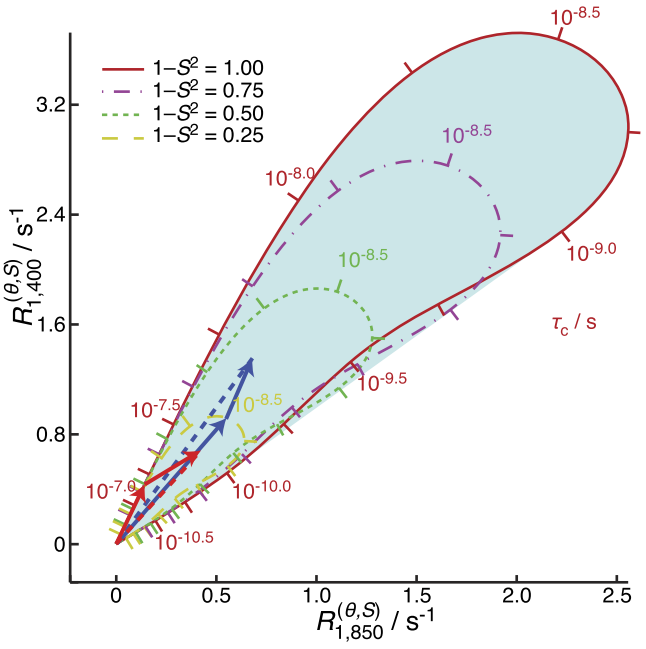

FIG. 2. Allowed region for ^{15}N R_1 acquired at 400 MHz and 850 MHz. All possible combinations for $R_{1,400}^{(\theta,S)}$ and $R_{1,850}^{(\theta,S)}$ are highlighted in light blue. Colored lines trace $R_{1,400}^{(\theta,S)}$ and $R_{1,850}^{(\theta,S)}$ for different S^2 values as a function of τ_c, for a mono-exponential correlation function. Colored arrows show how multiple correlation times (represented by blue or red solid arrows) add together and may be fitted with a single correlation time (represented by the dashed lines in the same color).

By definition, distributions of correlation times $(1-S^2)\theta(z)$ with more than one non-zero point (multi-exponential correlation functions) must also fall in the allowed range of pairs of relaxation-rate constants. Since the contributions of different correlation times to the relaxation-rate constants are additive [see Eqs. (5)–(9)], multi-exponential correlation functions are described by a point in the space of allowed rate constants that is the vector sum of the positions in the space for each individual correlation time and amplitude. Two examples for a bi-exponential correlation function are shown in Fig. 2. The first example considers a motion characterized by two correlation times, the first with a value of $\tau_c = 10^{-8.5}$ s, and corresponding amplitude of $(1-S^2)\theta(-8.5)dz = 0.25$, and a second with $\tau_c = 10^{-7.5}$ s and $(1-S^2)\theta(-7.5)dz = 0.5$ (so that $1-S^2 = 0.75$). The blue vectors indicate the contributions of the two motions, with the longer vector corresponding to $\tau_c = 10^{-8.5}$ s. The vector sum of the two motions (blue dashed line) corresponds to a final position in the plot that is also characteristic for a motion with a single exponential correlation function with $\tau_c = 10^{-8.41}$ s and $(1-S^2)\theta(-8.41)dz = 0.34$. This illustrates the ambiguity of relaxation data that can be the result of very different motional models and the fact that too simple models lead to results that have a bias towards the sensitive correlation times. Similarly, the red vectors illustrate a case for two motions, with $\tau_c = 10^{-7.5}$ and 10^{-10} s and $(1-S^2)\theta(z)dz = 0.5$ for the two exponentials (so that $1-S^2 = 1$). This point in space also corresponds to a motion with a single correlation time of $\tau_c = 10^{-8.63}$ s and an amplitude of $(1-S^2)\theta(-8.63)dz = 0.14$. In this case, R_1 is similarly sensitive to the two correlation times and the simple one-correlation time model leads to a correlation time that is roughly in the middle of the two correlation times.

The space of possible relaxation rate constants, for a given set of measurements, helps one to visualize the information available to describe the distribution of motion, $(1-S^2)\theta(z)$, including the ambiguity in the experimental data as discussed above. One may also use this space to optimize the separation of information for different ranges of correlation times. As shown in Fig. 1, it is possible to take a linear combination of the sensitivities of R_1 rate constants ($R_{1,\zeta}(z)$) to obtain two detectors with optimally separated sensitivities without either sensitivity becoming negative. This linear combination, defined in Eq. (10), is basically a coordinate transformation from the black coordinate system in Fig. 3(a) to the coordinate system defined by the red and blue axes [Figs. 3(a) and 3(b)] which are vectors tangential to the allowed region (blue area). We see that requiring the detector sensitivities to be optimally separated but non-negative results in a modified allowed region for the detector responses, $\rho_n^{(\theta,S)}$ [Fig. 3(b)], which spans most of the range for which $0 \le \rho_1^{(\theta,S)} \le 1$ and $0 \le \rho_2^{(\theta,S)} \le 1$, i.e., the first quadrant for a two-dimensional problem, but does not become negative for either $\rho_n^{(\theta,S)}$.

The coordinate system in Fig. 3(b) is spanned by the blue and red unit vectors $\vec{e}_1 = (1,0)^T$ and $\vec{e}_2 = (0,1)^T$, respectively. These vectors can be transformed back from the space of detector responses into the space of relaxation-rate

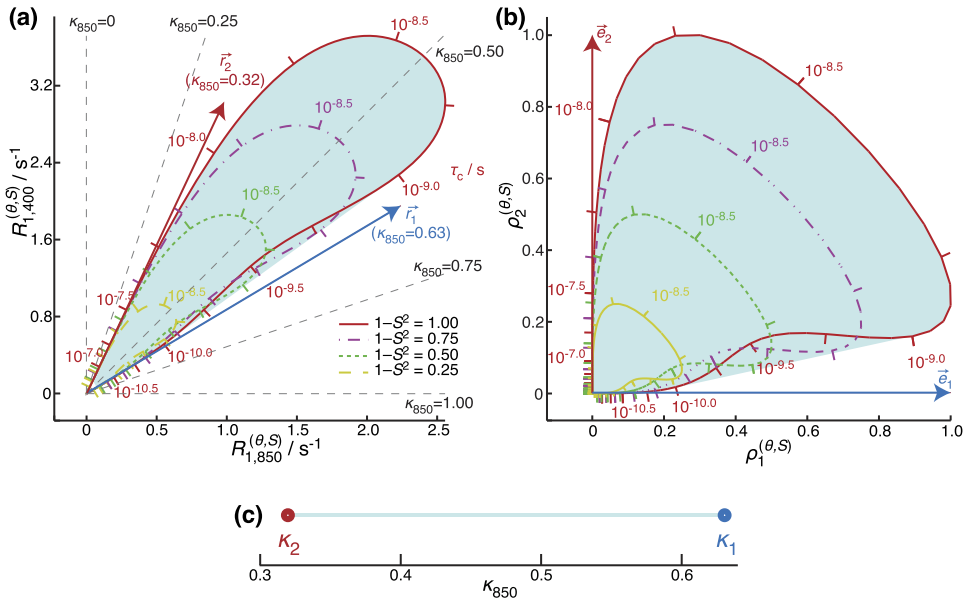

FIG. 3. Linear transformation of the allowed space and selection of detection vectors. (a) shows the allowed region of ^{15}N rate constants in blue (for 400, 850 MHz ^{1}H frequency), with κ_{850} values plotted as gray, dashed lines through the space. (b) shows the transformation of the allowed region, defined by Eq. (10), resulting in the allowed region of detector responses, $\rho_n^{(\theta,S)}$. Red and blue detection vectors in (a) and (b) are tangential to the space, and after the transformation in (b), these become unit vectors in the x and y directions ($\vec{e}_1, \vec{e}_2$). (c) is the 1D reduced space, defined by κ_{850}, where the allowed region is again shown in blue. The red and blue circles in the reduced space can be used to determine the optimal direction (although not the length) of the detection vectors.

constants, by inverting Eq. (10), with the resulting vectors ($\vec{r}_1$, $\vec{r}_2$) shown in Fig. 3(a). We refer to $\vec{r}_n$ as *detection vectors*. Here, we have obtained the detection vectors by inverting the transformation from the space of relaxation-rate constants to the space of detector responses. However, we can also do the reverse: place vectors in the space of relaxation rate constants, in order to define the transformation from relaxation rate constants to detector responses. This can be done as follows, resulting in values for *a*-*d*:

$$\begin{aligned}\begin{pmatrix}\vec{r}_1 & \vec{r}_2\end{pmatrix} &= \begin{pmatrix}a & b\\ c & d\end{pmatrix}^{-1}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},\\ \begin{pmatrix}a & b\\ c & d\end{pmatrix} &= \begin{pmatrix}\vec{r}_1 & \vec{r}_2\end{pmatrix}^{-1}.\end{aligned} \tag{13}$$

An alternative way to understand this transformation is to note that the detector responses can be considered to be weights such that the observed relaxation rate constants are given by a linear combination of the detection vectors,

$$\begin{pmatrix}R_{1,400}^{(\theta,S)}\\ R_{1,850}^{(\theta,S)}\end{pmatrix} = \begin{pmatrix}\vec{r}_1 & \vec{r}_2\end{pmatrix}\begin{pmatrix}\rho_1^{(\theta,S)}\\ \rho_2^{(\theta,S)}\end{pmatrix}. \tag{14}$$

Inversion leads to the following formula for obtaining detector responses from experimental data:

$$\begin{pmatrix}\rho_1^{(\theta,S)}\\ \rho_2^{(\theta,S)}\end{pmatrix} = \begin{pmatrix}\vec{r}_1 & \vec{r}_2\end{pmatrix}^{-1}\begin{pmatrix}R_{1,400}^{(\theta,S)}\\ R_{1,850}^{(\theta,S)}\end{pmatrix}. \tag{15}$$

Similarly, the detection vectors may be used to obtain the sensitivities of each detector,

$$\begin{pmatrix}\rho_1(z)\\ \rho_2(z)\end{pmatrix} = \begin{pmatrix}\vec{r}_1 & \vec{r}_2\end{pmatrix}^{-1}\begin{pmatrix}R_{1,400}(z)\\ R_{1,850}(z)\end{pmatrix}. \tag{16}$$

If we take the approach of placing detection vectors in the space of relaxation-rate constants, then a good strategy is to place them tangential to the allowed region. Then, the transformation leads to the stretched space of Fig. 3(b) that covers as much of the range as possible for which $0 \le \rho_1^{(\theta,S)} \le 1$ and $0 \le \rho_2^{(\theta,S)} \le 1$, without becoming negative. This leads to well-separated and non-negative detector sensitivities, $\rho_n(z)$, as discussed above. Note that in some cases, we may want to define our detectors such that it is possible to have $\rho_n^{(\theta,S)} > 1$, where the maximum value of $\rho_n^{(\theta,S)}$ is determined by the detection vector lengths.

In the case of two relaxation-rate constants, it is relatively straightforward to place the two detection vectors tangentially. As we add rate constants, however, the problem becomes higher-dimensional so that it becomes helpful to be able to visualize at least three rate constants in a single two-dimensional (2D) plot or four rate constants in a three-dimensional (3D) plot. Therefore, in order to reduce the dimensionality of the problem, we introduce the reduced space of rate constants.

The first step in defining the reduced space is to normalize the rate constants by their maximum possible value, arriving at the normalized rate constant, $\Re_\zeta^{(\theta,S)}$,

$$\begin{aligned}\Re_\zeta^{(\theta,S)} &= \frac{R_\zeta^{(\theta,S)}}{c_\zeta},\\ c_\zeta &= \max(R_\zeta(z)).\end{aligned} \tag{17}$$

This step is not critical in the case of rate constants that have similar maximum values, but we will always include it for consistency. It is convenient for subsequent visualization of the reduced space, for example, when one combines R_1 and $R_{1\rho}$ measurements in a dynamics analysis since the maximum possible $R_{1\rho}$ rate constants are potentially several orders of magnitude larger than those for R_1. The second step is to divide all normalized rate constants by the sum of all normalized rate constants such that one obtains

$$\begin{aligned}\kappa_{850} &= \frac{\Re_{1,850}^{(\theta,S)}}{\Sigma_\zeta \Re_\zeta^{(\theta,S)}},\\ \Sigma_\zeta \Re_\zeta^{(\theta,S)} &= \Re_{1,400}^{(\theta,S)} + \Re_{1,850}^{(\theta,S)}.\end{aligned} \tag{18}$$

We refer to the result, κ_{850}, as the ratio of normalized rate constants (later shortened to the ratio of rates).

We could just as well define the space with κ_{400} (given by $\Re_{1,400}^{(\theta,S)}/\Sigma_\zeta \Re_\zeta^{(\theta,S)}$), but using both is redundant because we define them in such a way that they are not independent since their sum is always one. The divisor is chosen to factor out the absolute size of the rate constants so that the information is stored in their ratios. In other words, we factor out the dependence on $1 - S^2$ and only retain information on the relative distribution over the correlation times, $\theta(z)$. We want to avoid having the divisor approach zero faster than the numerator so that the sum of all normalized rate constants is a good choice for any set of relaxation rate constants in solid-state NMR. Note that in the case of three or more relaxation rate constants, there will be multiple κ_ζ to specify the location in the reduced space. Each ratio of rates, κ_ζ, must then be divided by the same $\Sigma_\zeta \Re_\zeta^{(\theta,S)}$, and the set of κ_ζ values will be denoted as a vector, $\vec{\kappa}$, which we refer to as the *reduced vector*. In Fig. 3(a), several values of κ_{850} are indicated in the space (dotted lines), where we can see that κ_{850} defines the ratio of the two rate constants, but not their absolute magnitude. Figure 3(c) then plots the reduced space, now a one-dimensional (1D) space, with the allowed region of the reduced space highlighted in blue.

In the reduced space, defined by κ_{850}, it becomes simple to select vectors tangential to the space. One simply places a point at the maximum and minimum values in the allowed region. These points can be used to define the direction of the detection vectors as follows:

$$\vec{r}_n = a_n \begin{pmatrix} \kappa_{400} c_{400} \\ \kappa_{850} c_{850} \end{pmatrix}, \qquad \kappa_{400} = 1 - \kappa_{850}. \tag{19}$$

First, note that the length is not defined by κ_{850} such that a_n must be determined separately. The length of each detection vector, $\vec{r}_n$, is inversely proportional to the magnitude of the sensitivity of that detector, $\rho_n(z)$ [see Eq. (14)]. For the moment, we will choose a_n so that $\rho_n(z)$ has a maximum of one (normalization will be discussed in Sec. III G). Second, we note that one of the κ_ζ must always be calculated from the other κ_ζ values since the dimensionality of the space has been reduced by 1, but this is straightforward since all κ_ζ sum to 1.

When one includes more relaxation-rate constants, ideally one wants to place reduced vectors in the reduced space that exactly surround the allowed region. However, this is only possible if the allowed region forms a polytope with exactly N corners, where N is the number of experiments included. For a one-dimensional reduced space, one always obtains a line so that we have two "corners," but for more experiments, only special cases will result in such a space (for example, a triangle for three experiments and a tetrahedron for four experiments). The different cases will be further discussed in the following sections.

First, however, we want to better understand how detector responses are related to distributions of motion. To understand what information we obtain from the detectors, we consider several distributions of motion, plotted in Fig. 4(a) (solid lines),

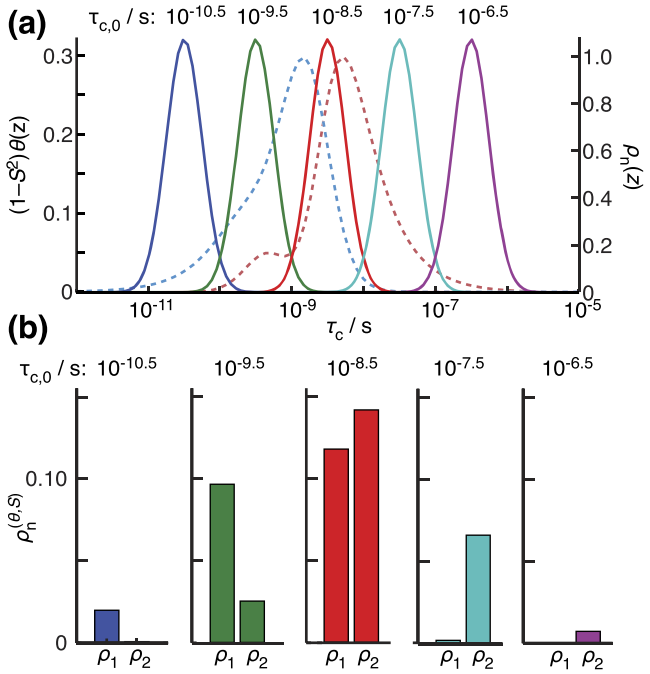

FIG. 4. Response of $\rho_n^{(\theta,S)}$ detectors to distributions of motion. (a) shows five distributions of motion, $(1 - S^2)\theta(z)$ (solid lines, left axis), and their overlap with the sensitivities of two detectors, $\rho_n(z)$ (dashed lines, right axis). The distributions of motion are log-Gaussian distributions, with $1 - S^2 = 0.2$ and $\sigma = 0.25$, and the center of each distribution $\tau_{c,0}$ indicated above each distribution.[19,28] (b) shows the response of the two detectors, $\rho_n^{(\theta,S)}$, to each distribution of motion given by Eq. (12). The center of each distribution ($\tau_{c,0}$) is indicated above each Gaussian and bar plot. The distributions (a) and resulting detector responses (b) are given in the same color, and the responses are positioned approximately below the corresponding distribution.

each taking the form of a log-Gaussian distribution,[28]

$$\theta(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z - z_0)^2}{2\sigma^2}\right). \tag{20}$$

The detector sensitivities are also plotted (dashed lines). Then, the overlap of the distribution ($(1 - S^2)\theta(z)$) and the sensitivity ($\rho_n(z)$) give the response of each detector, $\rho_n^{(\theta,S)}$, as shown in Fig. 4(b) [see Eq. (12)]. One observes the following behavior: if most of the motion has correlation times shorter than $\sim 10^{-10}$ s, then fitting of experimental data yields nonzero values for $\rho_1^{(\theta,S)}$ but approximately zero for $\rho_2^{(\theta,S)}$ [Fig. 4(b), left]. Conversely, if most of the motion has correlation times longer than 10^{-8} s, then $\rho_1^{(\theta,S)}$ is approximately zero, but $\rho_2^{(\theta,S)}$ is not [Fig. 4(b), right]. If the motion falls in between the detector sensitivities, then both detectors are non-zero [Fig. 4(b), center]. This is, however, the same response that one obtains if there are larger motions with one correlation time shorter than 10^{-10} s and one longer than 10^{-8} s so that these cases are not distinguishable with only two relaxation-rate measurements. Then, with several detectors, one can compare motion in different ranges of correlation times and can compare motion within the sensitivity range of each detector to other residues. Without further assumptions, one cannot get quantitative measurement of the motional amplitudes. However, it is possible to the get the average value of $1 - S^2$ under certain conditions or to estimate the amplitude of the distribution function, $(1 - S^2)\theta(z)$, at particular correlation times (see Sec. III G).

B. Three relaxation rate constants

In Sec. III A, a method for visualizing the ambiguity of a set of two relaxation measurements was presented, and a procedure for fitting experimental data without assumptions about the complexity of the correlation function was demonstrated. This approach needs to be generalized for larger data sets with more than two relaxation measurements. To demonstrate the general approach, we add an additional relaxation-rate constant, ^{15}N R_1, measured at 600 MHz.

The full space of normalized rate constants is shown as a 3D plot in Fig. 5(a). Now, the allowed region is represented as a volume, albeit with a rather thin extension in one direction. Red and blue traces show positions in the space resulting from a mono-exponential correlation time as a function of τ_c, for two values of $1 - S^2$. One sees that distributions having only a mono-exponential correlation function no longer span all points in the allowed region. The thin extension essentially tells us that the additional relaxation measurement did not significantly increase our ability to further characterize the distribution of motion, $(1 - S^2)\theta(z)$. With this in mind, we discuss the procedure to determine the detectors, $\rho_n^{(\theta,S)}$, as we did with two relaxation-rate constants.

As before, we use a representation that removes the dependence of the rate constants on $1 - S^2$. Here, we use the axes κ_{850} and κ_{600} defined as

$$\kappa_{850} = \frac{\Re_{1,850}^{(\theta,S)}}{\Sigma_\zeta \Re_\zeta^{(\theta,S)}},$$
$$\kappa_{600} = \frac{\Re_{1,600}^{(\theta,S)}}{\Sigma_\zeta \Re_\zeta^{(\theta,S)}}, \tag{21}$$
$$\Sigma_\zeta \Re_\zeta^{(\theta,S)} = \Re_{1,400}^{(\theta,S)} + \Re_{1,600}^{(\theta,S)} + \Re_{1,850}^{(\theta,S)}.$$

This two-dimensional plot of the allowed space is shown in Fig. 5(b). Note that now a position in the space is represented by two κ values, given by the reduced vector, $\vec{\kappa}_n$, representing a set of ratios of all relaxation rate constants. If we plot the positions in the space for a mono-exponential correlation function as a function of the correlation time, τ_c, we can see that this trace only covers the upper border of the allowed region of the reduced space (plus a loop through the space). This is in contrast to the discussion in Sec. III A with two relaxation rate constants where all points in the reduced space can be characterized by a mono-exponential correlation function. For three relaxation rate constants, we need two correlation times to cover the allowed region completely. The trace of mono-exponential correlation functions is color coded with the value of $\Sigma_\zeta \Re_\zeta^{(\theta,S)}$, where $1 - S^2 = 1$ (denoted $\Sigma_\zeta \Re_\zeta^{(\theta,0)}$).

If we want to fit three relaxation rate constants, the most convenient way is by picking values for three reduced vectors, $\vec{\kappa}_n$, and calculating the corresponding detection vectors, $\vec{r}_n$. Figure 6(a) shows the result of picking three reduced vectors within the allowed region of the reduced space, the fit of the relaxation rate constants for mono-exponential correlation functions, and the resulting sensitivities, $\rho_n(z)$. As will always be the case if we have three rate constants and three fit parameters, the rate constants are exactly fit unless the three $\vec{\kappa}_n$ are not linearly independent. The solution of Fig. 6(a) corresponds to one of the solutions where the three vectors are tangential to the allowed space. However, we note that in contrast to the example with two relaxation rate constants, the triangle of the three reduced vectors does not surround the complete allowed region leading to negative values for one or more of $\rho_n(z)$. When fitting a distribution of motion, this is not ideal because one correlation time in the distribution may yield a positive contribution to one of the responses, $\rho_n^{(\theta,S)}$, whereas that contribution can be partially or fully canceled by a negative contribution at another correlation time. An alternate solution with only positive sensitivities is obtained by placing the three reduced vectors further outside of the allowed region, as seen in Fig. 6(b), so that the complete allowed region is within the resulting triangle. There are many possible such solutions, but

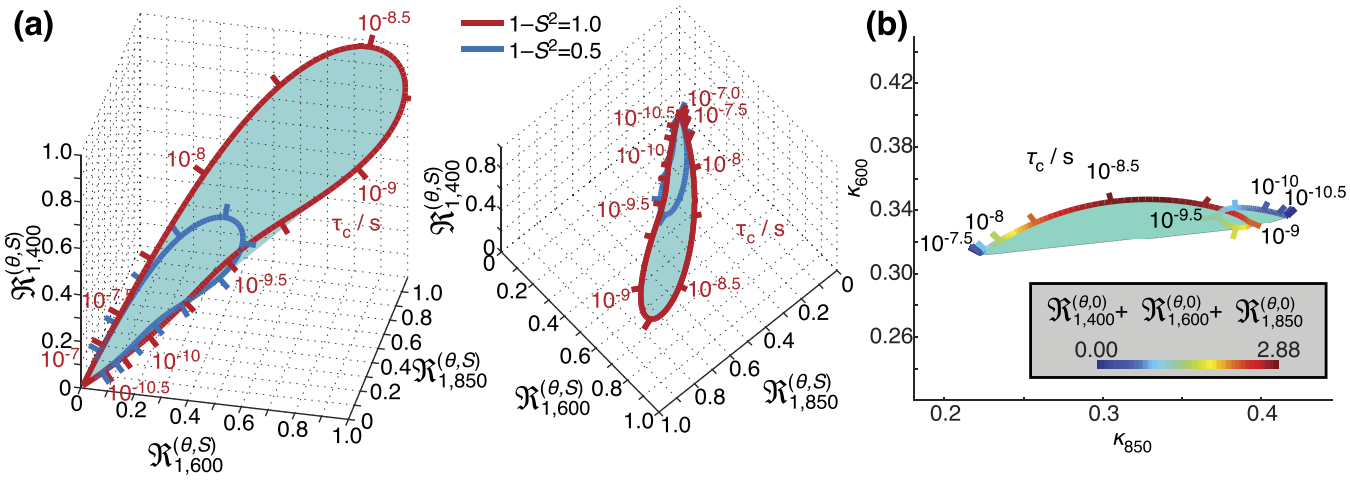

FIG. 5. Observable values for ^{15}N R_1 acquired at external fields of 400, 600, and 850 MHz. (a) shows two views of the 3D space, where blue regions highlight possible combinations of $\Re_{1,400}^{(\theta,S)}$, $\Re_{1,600}^{(\theta,S)}$, and $\Re_{1,850}^{(\theta,S)}$. Colored lines trace $\Re_{1,400}^{(\theta,S)}$, $\Re_{1,600}^{(\theta,S)}$, and $\Re_{1,850}^{(\theta,S)}$ for different S^2 values as a function of τ_c (τ_c runs from $-\infty$ to $+\infty$, although ticks are not shown beyond 10^{-11} and 10^{-6} s). (b) shows physically possible values of κ_{850} and κ_{600} in blue, and plots τ_c against κ_{850} and κ_{600}, while encoding $\Sigma_\zeta \Re_\zeta^{(\theta,S)}$ for each τ_c with color, assuming $1 - S^2 = 1$ (blue: slow relaxation and red: fast relaxation).

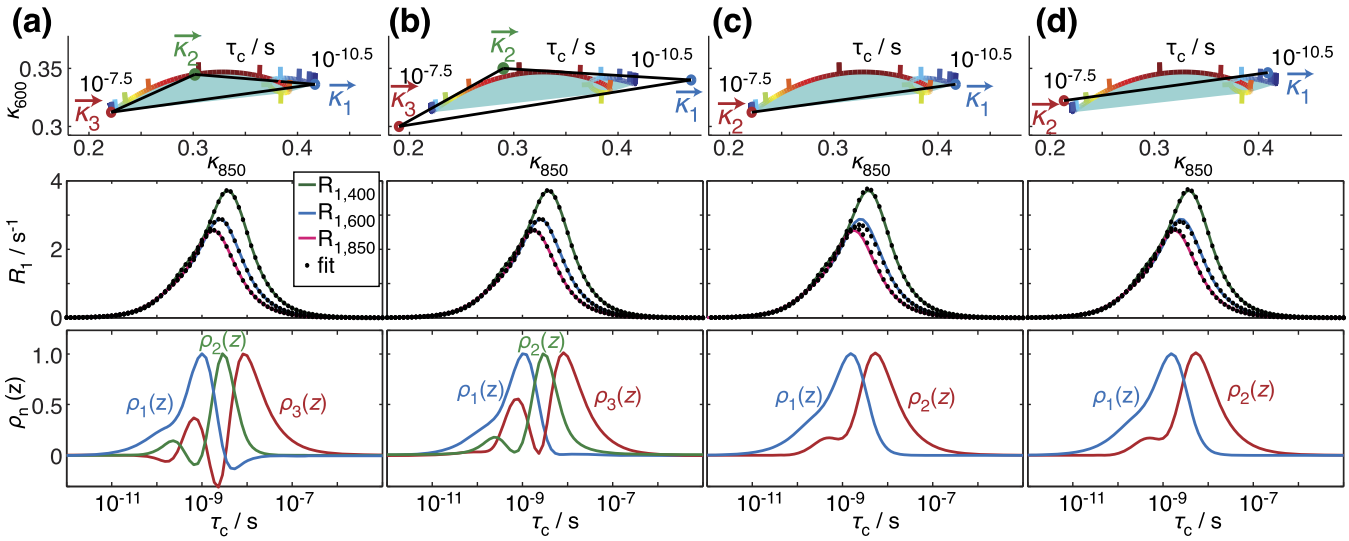

FIG. 6. Fitting options for measurement of ^{15}N R_1 at 400, 600, and 850 MHz. Four methods of fitting are shown: three $\vec{\kappa}_n$ within the allowed region (a), three $\vec{\kappa}_n$ surrounding the allowed region (b), two $\vec{\kappa}_n$ within the allowed region (c), and two $\vec{\kappa}_n$ outside the allowed region to improve fitting (d). In (a)-(d) (top), the 2D space defined by κ_{850} and κ_{600} is shown, with positions of the reduced vectors, $\vec{\kappa}_n$, shown (colored circles). The middle plot shows the R_1 rate constants resulting from a single correlation time, with $1 - S_2 = 1$ ($R_\zeta(z)$, colored lines), and the fit quality using detection vectors, $\vec{r}_n$, corresponding to $\vec{\kappa}_n$ shown at top (black dots). The bottom plot shows the sensitivities, $\rho_n(z)$, for each detection vector, $\vec{r}_n$.

we require the points to be placed as close as possible around the allowed region minimizing the forbidden area inside the possible positive linear combinations. This is similar to the two relaxation-rate example where the tangential vectors ensure that the complete allowed area can be represented by positive linear combinations while at the same time minimizing the forbidden area inside the space of positive linear combinations. In Fig. 6(b), the three reduced vectors were placed at the ends of the space, but we can no longer do this perfectly with three rate constants unless the space is perfectly triangular. In this case, measured rate constants that have shifted outside of the allowed region due to experimental noise will be forced into the space, therefore reducing noise in the fit parameters—assuming that a fit is performed using the criteria that all $\rho_n^{(\theta,S)} \geq 0$. Second, the functions $\rho_n(z)$ are better separated from each other if the reduced vectors are near the extrema of the allowed region.

The allowed region in the reduced space of Fig. 5(b) is rather narrow and can reasonably well be covered by selecting only two reduced vectors, $\vec{\kappa}_n$. Such a selection of detectors forces the fit of the experimental rate constants to collapse onto a line crossing through the allowed region. The fit of the experimental data is no longer perfect, but it would not lead to large deviations in the data fitting and would not require negative intensities. This is shown in Fig. 6(c) where reduced vectors have been placed at the two ends of the allowed region, although one notes that near $\tau_c = 10^{-8.5}$ s, there are some deviations visible in the fit of the relaxation rate constants. To improve this, the positions of $\vec{\kappa}_n$ can be shifted slightly upward, to better cover the middle of the allowed region. In this case, shown in Fig. 6(d), the fit of all rate constants is very good. Only with very high signal-to-noise experiments would it be possible to distinguish the fits of rate constants in Fig. 6(a)/6(b) from those in Fig. 6(d). Therefore, we conclude that for such a data set, one should usually only use two reduced vectors, otherwise the errors of the resulting detector responses, $\rho_n^{(\theta,S)}$, could be high. In practice, it may be helpful to try fitting more or fewer vectors to experimental data, to determine if over-fitting or under-fitting occurs via statistical analysis of the fitted data, using a method such as reduced-χ^2 or bias-corrected versions of the Akaike information criterion (AIC).[29–31] Note that acquiring more than two R_1 rate constants still provides considerable value to experimental data as it is a good means of verifying the data quality and improves the overall experimental signal to noise, although it will not usually add much information that could not be extracted from the first two rate constants alone. Higher or lower fields than the example here with 400 and 850 MHz may also allow additional detection vectors, as would more (4-5) R_1 measurements, or the inclusion of different nuclei, for example, backbone ^{15}N relaxation with ^{13}CO relaxation[13] and the inclusion of other experiments such as Nuclear Overhauser Effect (NOE) measurement.[32]

C. Different sensitivity ranges: Longitudinal and transverse relaxation

Now that we have described the analysis of multiple rate constants using multiple detectors to characterize distributions of motions, we consider a few categories of data sets. First, we consider what happens when we combine longitudinal (R_1) and transverse relaxation ($R_{1\rho}$) data in the case of solid-state NMR. We take as an example two ^{15}N R_1 measurements at 400 and 850 MHz and two $R_{1\rho}$ measurements at an external field of 850 MHz, MAS of 60 kHz, and spin-lock strengths of 15 and 45 kHz. The resulting space is plotted in Fig. 7(a), where only the reduced space is shown, since the 4D space cannot be plotted. In this case, selecting $\vec{\kappa}_n$ is straightforward since the allowed region is broad and nearly a tetrahedron. This is a result of the nearly separate sensitivity ranges of the R_1 and $R_{1\rho}$ rate constants [Fig. 7(b)]. Therefore, one simply places

one $\vec{\kappa}_n$ near each corner of the tetrahedron. The resulting fit of rate constants and sensitivities, $\rho_n(z)$, are plotted in Figs. 7(b) and 7(c).

By selecting $\vec{\kappa}_n$ only slightly away from the corners of the tetrahedron, one can force the matrix of $\vec{r}_n$ to be exactly block-diagonal,

$$\left(\vec{r}_1\ \vec{r}_2\ \vec{r}_3\ \vec{r}_4\right) = \begin{pmatrix} c_{1,400}\kappa^{(1)}_{R1,400} & c_{1,400}\kappa^{(2)}_{R1,400} & 0 & 0 \\ c_{1,850}\kappa^{(1)}_{R1,850} & c_{1,850}\kappa^{(2)}_{R1,850} & 0 & 0 \\ 0 & 0 & c_{1\rho,15}\kappa^{(3)}_{R1\rho,15} & c_{1\rho,15}\kappa^{(4)}_{R1\rho,15} \\ 0 & 0 & c_{1\rho,45}\kappa^{(3)}_{R1\rho,45} & c_{1\rho,45}\kappa^{(4)}_{R1\rho,45} \end{pmatrix}. \tag{22}$$

In this case, we see that there is almost no advantage to fitting the R_1 and $R_{1\rho}$ rate constants simultaneously since the block diagonal matrix shows that $\rho_1^{(\theta,S)}$ and $\rho_2^{(\theta,S)}$ will be fit only to the R_1 data, and $\rho_3^{(\theta,S)}$ and $\rho_4^{(\theta,S)}$ will only be fit to the $R_{1\rho}$ data. This is not too surprising, seeing that the sensitive ranges of R_1 and $R_{1\rho}$ have only a small overlap, as shown in Fig. 7(b), so that information from these two data types is nearly independent. It is worth noting, however, that correlation between dynamics obtained with R_1 and with $R_{1\rho}$ can imply either a broad distribution of correlation times or motion in between the R_1 and $R_{1\rho}$ sensitive ranges. Note that a number of studies have found similar trends in longitudinal relaxation (R_1) and transverse relaxation ($R_{1\rho}$ or CSA-dipole cross correlated relaxation) and have, therefore, successfully combined those data sets and fit with a single relaxation model, which implies that one of these cases is likely.[6,7,10,13,33,34]

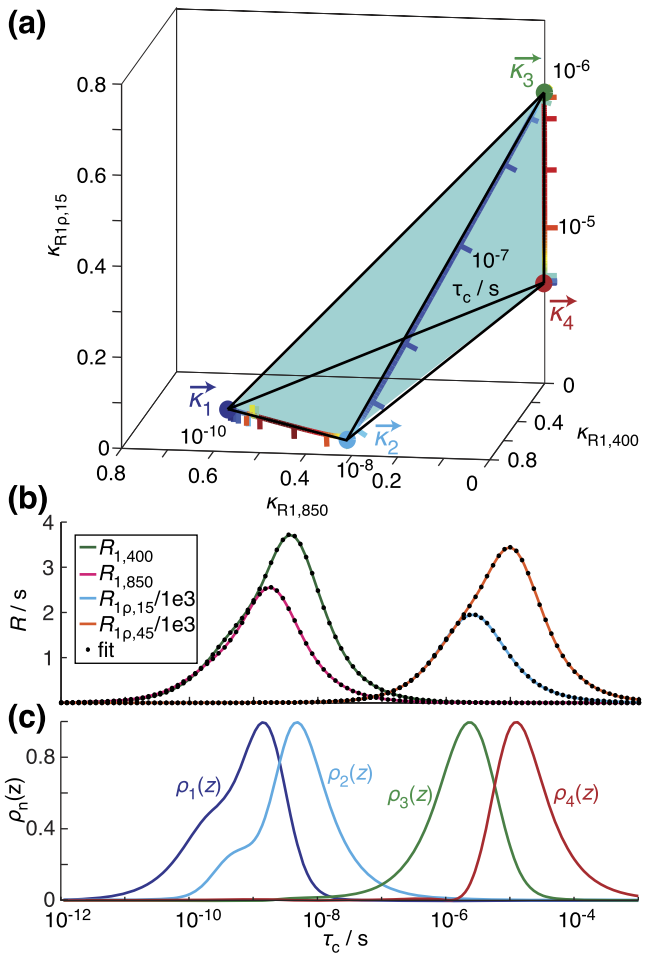

FIG. 7. Fitting of four measured rate constants (^{15}N R_1 at 400 and 850 MHz, $R_{1\rho}$ at a field of 850 MHz, MAS of 60 kHz, and field strength of 15 and 25 kHz). (a) shows the positions of the reduced vectors, $\vec{\kappa}_n$, in the space (colored circles), and black lines surround the allowed region that can be fitted with positive $\rho_n^{(\theta,S)}$. (b) shows the fitting of the measured rate constants using these $\vec{\kappa}_n$. (c) shows the sensitivities, $\rho_n(z)$.

D. Relation to spectral-density mapping (and related methods)

Spectral-density mapping is a strategy for NMR dynamics analysis that takes a linear combination of the relaxation-rate constants in order to calculate the value of the spectral-density functions at a few selected frequencies.[24] Since it takes a linear combination of the rate constants, it is a special case of the detectors method presented here. For ^{15}N R_1, R_2, and σ_{NH} measured at a single static magnetic field, these combinations approximate with high accuracy the following spectral densities:

$$\begin{aligned} J(0) &\approx \frac{R_2 - 0.5R_1 - 0.454\sigma_{NH}}{3\delta^2 + 2(\omega_I\sigma_{zz})^2}, \\ J(\omega_N) &\approx \frac{R_1 - 1.249\sigma_{NH}}{3\delta^2/4 + (\omega_I\sigma_{zz})^2/2}, \\ J(0.870\omega_H) &\approx \frac{4\sigma_{NH}}{5\delta^2}. \end{aligned} \tag{23}$$

Therefore, $J(0)$, $J(\omega_{\mathrm{N}})$, and $J(0.870\ \omega_{\mathrm{H}})$ are proportional to detectors $\rho_1^{(\theta,S)}$, $\rho_2^{(\theta,S)}$, and $\rho_3^{(\theta,S)}$ obtained via

$$\begin{pmatrix} \rho_1^{(\theta,S)} \\ \rho_2^{(\theta,S)} \\ \rho_3^{(\theta,S)} \end{pmatrix} \propto \begin{pmatrix} 1 & 0.5 & -0.454 \\ 0 & 1 & -1.249 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R_2 \\ R_1 \\ \sigma_{NH} \end{pmatrix}. \tag{24}$$

This is analogous to Eq. (15). Similarly, the sensitivity of these detectors is approximately given by the functional form of the spectral density at the given frequencies,

$$\rho_1(z) \propto 10^{-z},$$
$$\rho_2(z) \propto \frac{10^{-z}}{1 + (\omega_N 10^{-z})^2}, \tag{25}$$
$$\rho_3(z) \propto \frac{10^{-z}}{1 + (0.870\omega_H 10^{-z})^2}.$$

The reduced space defined by the ratio of rates for R_1, R_2, and σ_{NH} acquired at 850 MHz is shown in Fig. 8(a). The positions of the reduced vectors, $\vec{\kappa}_n$, corresponding to the matrix in Eq. (24) are shown in Fig. 8(a) and the resulting sensitivities, which are approximately proportional to the spectral densities [Eq. (25)], are shown in Fig. 8(b). Therefore, we see that spectral density mapping is a special case of the detectors method introduced here. However, one also notes that the separation of the sensitivities is suboptimal. If the positions of $\vec{\kappa}_n$ are moved to fall at the corners of the nearly triangular allowed region, then new sensitivities are obtained, shown in Fig. 8(c), where the separation of the detectors is improved. Of course, these no longer correspond to the spectral densities, but they are optimum for the characterization of the distribution of motion, $(1 - S^2)\theta(z)$.

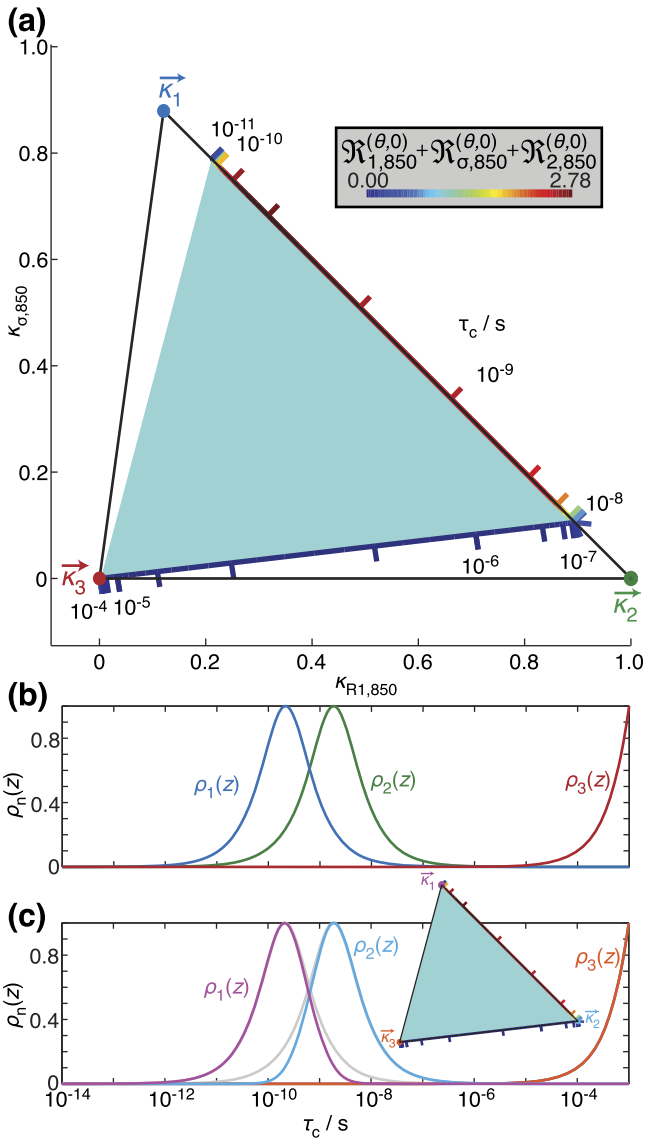

FIG. 8. Space defined by R_1, R_2, and σ_{NH} at 850 MHz. (a) shows the reduced space, along with positions of $\vec{\kappa}_n$ defined by the spectral density mapping approach. (b) shows the sensitivities of corresponding detectors, which are approximately proportional to $J(0)$, $J(\omega_N)$, and $J(0.870\omega_H)$. (c) shows sensitivities corresponding to reduced vectors placed at corners of the allowed region (see inset for positions of $\vec{\kappa}_n$), yielding better separation of the sensitivities (the spectral density mapping sensitivities are plotted in gray for comparison).

We also note that LeMaster used R_1, R_2, and NOE at the same magnetic field in a similar methodology,[35,36] in which he fitted the data to a correlation function of the form

$$C(t) = \frac{1}{5}\left[S^2 + S_f^2\left(S_H^2 S_N^2 \exp(-t/\tau_c) + (1 - S_H^2)\exp(-t/\tau_H) + S_H^2(1 - S_N^2)\exp(-t/\tau_N)\right)\right],$$
$$\tau_H = 1/(\omega_H + \omega_N), \tau_N = -1/\omega_N. \tag{26}$$

Here, τ_c is the rotational correlation time of the molecule in solution, and τ_H and τ_N are fixed where the experiments are sensitive. Fixing the correlation times results in a fit that is essentially a linear combination, as in the detectors approach. This approach has the advantage that it accounts for the overall tumbling in solution but is limited to a specific data set (R_1, R_2, NOE) and the sensitivities of each term in the correlation function are not explicitly defined. More recently, Ferrage and co-workers also introduced the IMPACT method of dynamics analysis,[37] where the correlation function was taken to have the form

$$C(t) = \frac{1}{5}\sum_{i=1}^{n} A_i \exp(-t/\tau_i),$$
$$\sum_{i=1}^{n} A_i = 1. \tag{27}$$

Again, the correlation times were fixed (the A_i are then fitted) so that they were logarithmically spaced. This contrasts with the LeMaster method, where the correlation times were specifically chosen to occur where experiments are sensitive, and contrasts with the detectors method where detection vectors are specifically placed in the reduced space.

E. Information content

When deciding on the number of required detection vectors, $\vec{r}_n$, and deciding on placing them, it is helpful to have an idea how much information is accessible from the experimental data. This will depend both on the set of experiments used and the signal-to-noise ratio of those experiments. To investigate how much information can be obtained from a set of experiments, we take three data sets and determine how accurately each data set may be fit as a function of the number of detection vectors. Accuracy is determined first by calculating and fitting $R_{1,\zeta}^{(\theta,S)}$ that result from a uniform distribution of motion and reporting the average over the fit error of each of the rate constants (the error is normalized by the magnitude of the resulting rate constant). A second measure of the accuracy is to calculate $R_{1,\zeta}^{(\theta,S)}$ for distributions of motions having only a single correlation time and determine the fit error, again reporting the average over the fit error for each of the rate constants. In the latter case, we report the largest error for all correlation times that were used.

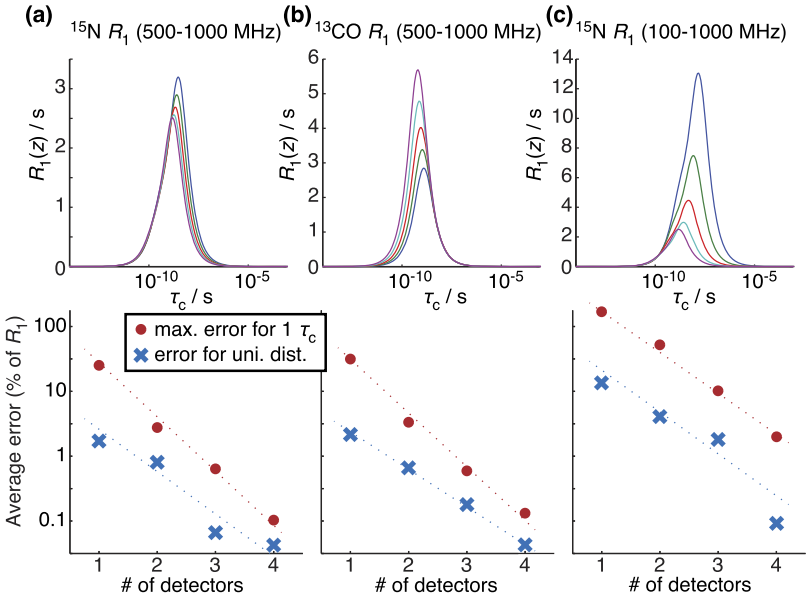

FIG. 9. Fitting error of relaxation rate constants as a function of the number of detectors used for fitting. The top plots show $R_1(z)$ for five rate constants. The bottom plots show the average error of the five rate constants, as a percentage of the rate constants. The average error is shown for a uniform distribution of motion (blue crosses) and the error resulting for a mono-exponential correlation function, for which τ_c is taken to be that which gives the maximum error. (a) uses five ^{15}N R_1 rate constants with magnetic fields distributed logarithmically between 500 and 1000 MHz. (b) uses 5 ^{13}CO rate constants distributed logarithmically between 500 and 1000 MHz. (c) uses five ^{15}N R_1 rate constants distributed logarithmically between 100 and 1000 MHz.

This is done for three sets of rate constants, the first for ^{15}N R_1 acquired at five magnetic fields, spaced logarithmically between 500 and 1000 MHz [Fig. 9(a)]. The second, we use ^{13}CO R_1, at the same fields [Fig. 9(b)]. This is done to test the importance of the width of the sensitivity of each R_1 rate constants, where ^{13}CO R_1 relaxation has a narrower sensitivity, $R_{1,\zeta}(z)$, because ^{13}CO relaxation is dominated by CSA relaxation and depends mostly on $J(\omega_I)$ [see Eq. (7)]. Finally, we use ^{15}N R_1 relaxation spaced logarithmically between 100 and 1000 MHz to test the importance of the spacing of the measured rate constants.

The results are shown in Fig. 9, bottom, where the averaged error is plotted on a logarithmic scale, as a percentage of the total rate constant. For a real distribution, we expect the average error to fall somewhere between the results for the mono-exponential and uniform distributions. Then, in Fig. 9(a), we see that we expect the error of the fit to be between ~2% and 25% if using a single detection vector, $\vec{r}_1$. Note that this is the average error so that some of the individual experiments will have higher error. Then, if experiments are accurate enough to distinguish this error, we can introduce a second detection vector. The requirements for experimental accuracy are exponentially increasing for each additional detector, as one sees that in the logarithmic plot in Fig. 9(a), the error decreases approximately linearly for each new detector. This trend is reproduced in Figs. 9(b) and 9(c). In Fig. 9(b), we see almost the same level of error so that there is no additional information resulting from the fact that we use a rate constant with a narrower sensitivity, $R_{1,\zeta}(z)$. In Fig. 9(c), we do increase the error for each number of detection vectors so that by having a broader range of magnetic fields for R_1 measurements, we are able to fit more detection vectors, thereby obtaining more information on the internal dynamics.

F. Scaled anisotropic interactions

Scaled anisotropic interactions characterized by an order parameter S^2 are used in solid-state NMR dynamics measurements to characterize the total amplitude of motion.[7,38] Because of their very broad range of sensitivity (~10 fs–~1 ms), they can be used here to give information about the amplitude of motion where the measured relaxation rate constants are not sensitive. We can define an additional detector, $\rho_0^{(\theta,S)}$, whose sensitivity is given by

$$\rho_0(z) = \frac{1}{a_0}\left(1 - \sum_{n=1} b_n \rho_n(z)\right). \tag{28}$$

The values of b_n and a_0 depend on the normalization scheme to be used, which is discussed below (Sec. III G). A simple case where a_0 and b_n are all set to 1 is shown in Fig. 10, with the same $\rho_n^{(\theta,S)}$ as were shown in Fig. 7(c). Note that to obtain $\rho_0^{(\theta,S)}$ from measured rate constants, one must modify the formula given in Eq. (16) to include fitting of $1 - S^2$ as follows:

$$\begin{pmatrix} \rho_0 \\ \rho_1 \\ \vdots \\ \rho_N \end{pmatrix} = \begin{pmatrix} \vec{0} & \vec{r}_1 & \cdots & \vec{r}_N \\ a_0 & b_1 & \cdots & b_n \end{pmatrix}^{-1} \begin{pmatrix} R_\zeta \\ \vdots \\ \vdots \\ 1 - S^2 \end{pmatrix}, \tag{29}$$

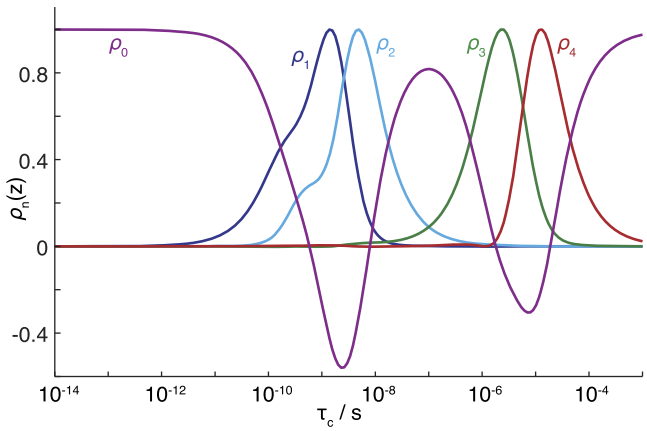

FIG. 10. Sensitivity, $\rho_0(z)$, when including an RDC measurement. Experiments are the same as those in Fig. 7, plus an additional direct measurement of $1 - S^2$.

where $\vec{r}_n$ are as defined before and $\vec{0}$ is a vector of zeros the same length as the $\vec{r}_n$. Possible definitions of a_0 and b_n are discussed in Sec. III G. Note that if $\rho_0(z)$ is negative at some points, then the response, $\rho_0^{(\theta,S)}$, may also be negative for some distributions of motion. The negative values of $\rho_0(z)$ occur at locations where other detectors overlap and will be discussed in more detail in Sec. III G.

G. Normalization schemes

So far, we have discussed how detection vectors, $\vec{r}_n$, may be selected to fit NMR relaxation data, and how this selection results in sensitivities, $\rho_n(z)$. The overlap of these sensitivities, with the distribution function, $(1 - S^2)\theta(z)$, yields the response of each detector, $\rho_n^{(\theta,S)}$ [Eq. (12)]. A residue specific detector response $\rho_n^{(\theta,S)}$ can be used to compare motions in different regions of the protein for a particular time scale of motion. However, there is also information contained in the relative responses of the different $\rho_n^{(\theta,S)}$ and in the absolute value of those responses; although without making assumptions about the motional model, it is not possible to obtain the exact amplitude of motion in a particular range of correlation times.

The normalization scheme we have used so far sets the maximum of each sensitivity, $\rho_n(z)$, to 1,

$$\max(\rho_n(z)) = 1. \tag{30}$$

We refer to this scheme as *equal-maximum* normalization. This scheme is a good choice for characterizing the contribution of each detector to the total amplitude of motion $(1 - S^2)$. First, we note that if direct measurement of $1 - S^2$ is included in the experimental data set, then the responses of all detectors add to yield $1 - S^2$, in other words,

$$\sum_{n=0} \rho_n^{(\theta,S)} = \sum_{n=0} (1 - S^2) \int \theta(z)\rho_n(z)dz = 1 - S^2, \tag{31}$$

which is the case because

$$\sum_{n=0} \rho_n(z) = 1 \tag{32}$$

for all z.

Furthermore, for a uniform distribution of motion, each detector response is then proportional to the *effective width* of that detector, with the width defined as the detector integral (area under curve) divided by its maximum,

$$\Delta z_n = \frac{\int_{-14}^{-3} \rho_n(z)dz}{\max(\rho_n(z))}. \tag{33}$$

Note that normalization of $\rho_0(z)$ is truncated at 10^{-14} s since it does not approach zero at short correlation times. This lower limit corresponds approximately to the frequency of the H–N stretch (3300 cm^{-1});[39] this vibration yields approximately the averaged H–N bond length of 1.02 Å,[40] from which the anisotropy of the ^1H–^{15}N dipole coupling is typically calculated ($\delta_{HN}/2\pi = -22.945$ kHz). The upper limit estimates where the Rotational Echo Double Resonance (REDOR)[50] experiment is no longer sensitive to motion, although its exact value depends on the experimental settings. The limits are not relevant for the other detectors since their sensitivities approach zero for both long and short correlation times.

In addition to calculating the effective width, Δz_n, for each detector, it is also useful to calculate the center of each detector, defined to be

$$z_n^0 = \frac{\int_{-14}^{-3} z\rho_n(z)dz}{\int_{-14}^{-3} \rho_n(z)dz}. \tag{34}$$

Then, for a detector analysis, the response of a detector, $\rho_n^{(\theta,S)}$, approximates its contribution to the total $1 - S^2$, around the correlation time given by z_n^0, for a range of correlation times given by Δz_n. Note this is an exact relationship for a uniform distribution of motion, but as the distribution becomes more irregular, then $\rho_n^{(\theta,S)}$ may deviate significantly from the total amplitude of motion in a range Δz_n around z_n^0.

A caveat in this approach is that in order for all detectors to have a maximum of one and for the sensitivities to sum to one, the sensitivity of the $\rho_0(z)$ detector must become negative at some points. This occurs where two (or more) of the other detector sensitivities strongly overlap, so they essentially detect some of the same motions twice. In order for the total amplitude of motion to then be correctly predicted (i.e., in order for $\rho_n^{(\theta,S)}$ to sum to $1 - S^2$), the sensitivity, $\rho_0(z)$, must compensate this with negative values at these points. This could be avoided by decreasing the amplitudes of the overlapping detector sensitivities, but then we would instead underestimate the amplitude of motion where the detectors do not overlap. Note that in special cases, the experimental $\rho_0^{(\theta,S)}$ may be negative, which then is a clear indication of motion in a region of overlapping sensitivities (often, negative contributions to $\rho_0^{(\theta,S)}$ are canceled by positive contributions so that we cannot easily identify such behavior).

The first approach tries to approximate the total amplitude of motion each detector sees, with its response both proportional to the amplitude of motion and the width of the detector. It is also possible, however, to directly estimate the amplitude of the distribution function, $(1 - S^2)\theta(z)$, at several correlation times ($z = \log_{10}(\tau_c)$). To do this, we first define a new normalization scheme, referred to as the *equal-integral* normalization, which requires for $\rho_n^{(\theta,S)}$,

$$\int_{-14}^{-3} \rho_n(z)dz = 1. \tag{35}$$

$\rho_0^{(\theta,S)}$ is also set to satisfy this integral by adjustment of a_0 in Eq. (28), but first we adjust b_n in Eq. (28) to minimize the sensitivity, $\rho_0(z)$, without any point in $\rho_0(z)$ going below zero. This allows optimal separation of the sensitivity of $\rho_0(z)$ from the other detectors, without yielding any negative sensitivities. The resulting sensitivities are shown in Fig. 11(b) (top).

This scheme is useful due to the following relationships. If we take the center of the sensitivity for each detector, z_n^0, then $\rho_n(z)$ are approximately symmetric around the corresponding z_n^0 (this is not the case for $\rho_0(z)$, which will be discussed below). Furthermore, if $\theta(z)$ is almost linear around z_n^0, then it may be approximated by $\theta(z) = \theta_0 + \theta_1 \times (z - z_n^0)$, where $\theta_0 = \theta(z_n^0)$ and θ_1 is the slope of the distribution function near z_n^0. In this

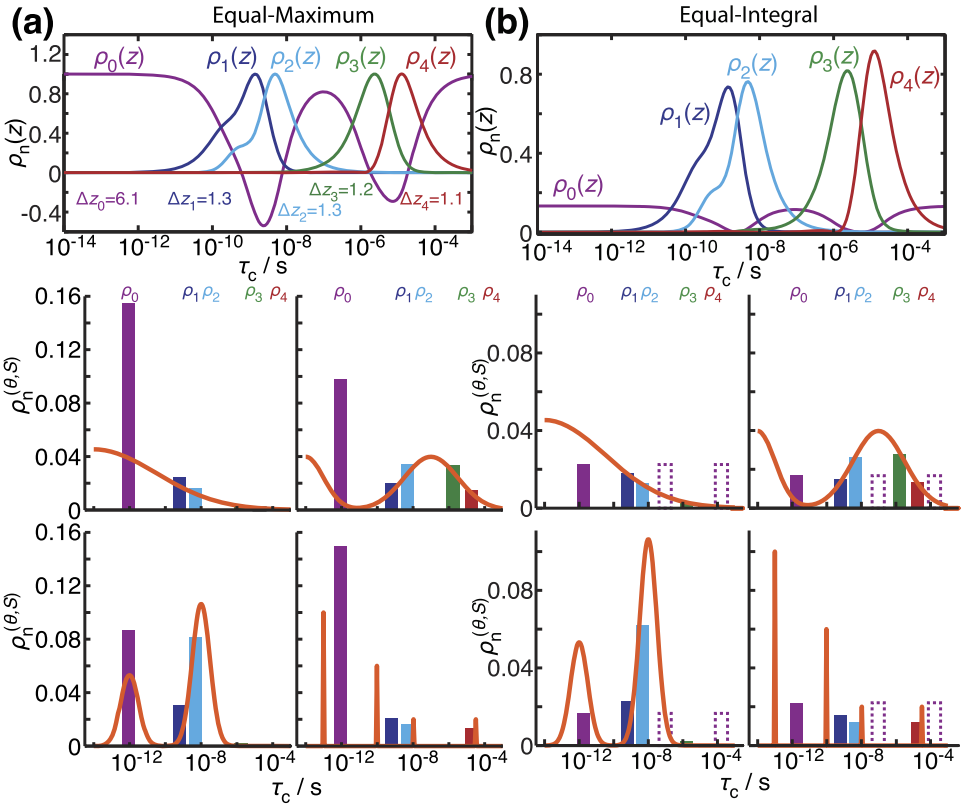

FIG. 11. Normalization schemes for detectors. The sets of experiments are the same as those shown in Fig. 7. The top plot in (a) shows $\rho_n(z)$, where $\rho_n(z)$ (excluding $\rho_0(z)$) has been normalized so that they have equal maxima, and their average response in the sensitive region is $1 - S^2$ (sensitive region highlighted in gray). The top plot in (b) shows $\rho_n(z)$ where all $\rho_n(z)$ have been normalized to have an integral of one. Lower plots show the response of ρ_n to several different distributions of motion, with the distribution of motion given as an orange line and bars giving the value of each ρ_n (color coded to match $\rho_n(z)$ at the top). Bars are placed at the center of the sensitivity range for each detector. Bars corresponding to $\rho_0(z)$ in (b) are also shown as dotted lines, where the sensitivity range has been split into three parts, and the detector response is shown at the center of each of the three ranges.

case, we may calculate the response of the detector, $\rho_n^{(\theta,S)}$, as follows:

$$\begin{aligned}\rho_n &= \left(1 - S^2\right) \int \theta(z)\rho_n(z)dz \\ &= \left(1 - S^2\right) \int \left(\theta_0 + \theta_1 \times (z - {z_n}^0)\right) \rho_n(z)dz \\ &= \left(1 - S^2\right) \theta_0 + \int \theta_1 \times (z - {z_n}^0)\rho_n(z)dz \\ &= \left(1 - S^2\right) \theta_0. \end{aligned} \tag{36}$$

The integral is zero in the last step since we assumed $\rho_n(z)$ is symmetric about z_n^0 and $\theta_1 \times (z - z_n^0)$ is anti-symmetric about z_n^0, thus yielding an overall anti-symmetric function which then integrates to zero. Then the amplitude of the distribution, $(1–S^2)\theta(z_n^0)$, may be estimated for distribution functions, where $\theta(z)$ is approximately linear (i.e., has a small second derivative). Relaxation rate constants are calculated for several distributions and fitted with $\rho_n^{(\theta,S)}$ shown in Fig. 11(b) (bottom). One sees that when the distribution function is approximately linear, $\rho_n^{(\theta,S)}$ estimate the distribution function $(1–S^2)\theta(z_n^0)$ very well. However, for the discrete distribution, this is no longer the case. Because $\rho_0(z)$ for solid-state NMR is not very symmetric about its center, z_n^0, we have split the function into three parts, separated by the positions where $\rho_0(z)$ is 0. For each region, a separate center (z_0) was calculated. Then, the resulting ρ_n is a weighted average of these three regions, with the weighting proportional to the width of each region [the two additional centers are shown as additional bars with dotted edges in Fig. 11(b)]. Accuracy is still subject to how symmetric $\rho_0(z)$ is in each region and how well $\theta(z)$ can be approximated by $\theta_0 + \theta_1 \times (z - z_n^0)$.

IV. EXAMPLE ANALYSIS WITH THE DETECTORS APPROACH

We have implemented the detector approach described so far in a numerical software package developed in MATLAB[41]

TABLE I. Detection vectors and parameters. *Other parameters:* $\delta_{HN}/2\pi$ = –22 945 Hz and $\Delta\sigma_N$ = 169.5 ppm.

	$\vec{r}_0$ (s^{-1})	$\vec{r}_1$ (s^{-1})	$\vec{r}_2$ (s^{-1})	$\vec{r}_3$ (s^{-1})	$\vec{r}_4$ (s^{-1})	B_0 (T)	$\omega_1/2\pi$ (kHz)	$\omega_r/2\pi$ (kHz)
$R_{1\rho,(13/110)}$	0	0	0	919.8	435.0	19.96	13	110
$R_{1\rho,(13/90)}$	0	0	0	1070.5	705.7	19.96	13	90
$R_{1\rho,(13/60)}$	0	0	0	956.0	1625.1	19.96	13	60
$R_{1,500}$	0	1.941	2.410	0	0	11.74	...	...
$R_{1,600}$	0	1.992	1.869	0	0	14.09	...	...
$R_{1,850}$	0	2.194	1.138	0	0	19.96	...	...
S^2	1.000	1.000	1.000	1.000	1.000	...	...	...

called DIstortion Free Relaxation Analysis TEchnique (DIFRATE),[42] distributed as open-source software under the terms of the GNU General Public License. In order to better understand the detectors approach, we present an example analysis based on solid-state NMR relaxation data previously published for ubiquitin dynamics using the DIFRATE software package. R_1 rate constants have been measured by Schanda *et al.*,[7] order parameters, S^2, have been measured by Haller and Schanda,[11] and $R_{1\rho}$ rate constants have been measured by Lakomek *et al.*[43] For each residue, the data set consists of up to three ^{15}N R_1 rate constants at fields of 500, 600, and 850 MHz, three $R_{1\rho}$ measurements acquired at 850 MHz, with a spin-lock strength of 13 kHz and MAS frequencies of 60, 90, and 110 kHz, and S^2 values acquired via REDOR.

We do not further discuss the selection of the detectors here, although a brief tutorial on detector optimization using the DIFRATE software is provided in the manual of the software package.[42] Note that for sake of reproducibility, it is important to clearly define the detectors used in an analysis, which can be done by providing the values found in the detection vectors. Furthermore, it is important to give all parameters required for calculating relaxation-rate constant sensitivity as a function of correlation time. These may be found in Table I. Detector sensitivity as a function of correlation time should usually be provided as a figure [see Fig. 12(a)], where exact values can be obtained from the information in Table I, using the DIFRATE software. Note that, as was discussed in Sec. III C, $\rho_1^{(\theta,S)}$ and $\rho_2^{(\theta,S)}$ are determined from R_1 rate constants and detectors $\rho_3^{(\theta,S)}$ and $\rho_4^{(\theta,S)}$ are determined from $R_{1\rho}$ rate constants. However, the matrix is not entirely block diagonal since $\rho_0^{(\theta,S)}$ depends on the values of all other detectors. Nonetheless, one may solve the problem for $\rho_1^{(\theta,S)}$, $\rho_2^{(\theta,S)}$ and then $\rho_3^{(\theta,S)}$, $\rho_4^{(\theta,S)}$ separately and later calculate $\rho_0^{(\theta,S)}$ according to the following equation:

$$\rho_0^{(\theta,S)} = \frac{1}{a_0}\left((1-S^2) - \sum_{n=1} b_n \rho_n^{(\theta,S)}\right), \tag{37}$$

which has a similar form as Eq. (28) (in this case, with $a_0 = 1$ and $b_n = 1$). Thus, one obtains a simplified problem with smaller matrices to solve.

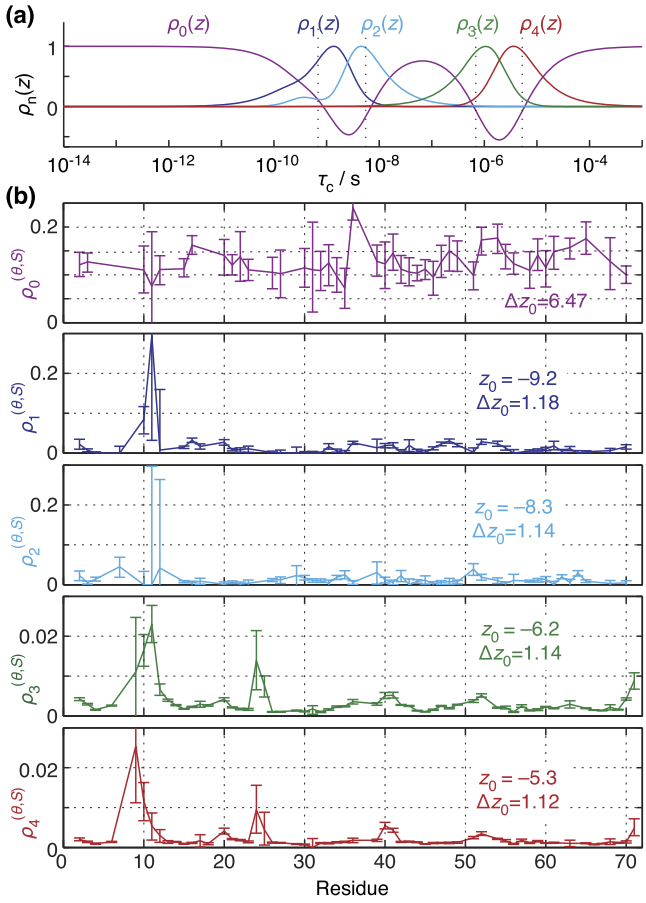

FIG. 12. Fit of detection vectors to ubiquitin dynamics data. Experimental data are taken from Refs. 7, 11, and 43. (a) shows the sensitivities of the five detectors, normalized according to the equal-max scheme, where the sensitive region is highlighted [defined in Eq. (12)]. (b) shows the residue specific response to each of the five detectors. Error bars give the 95% confidence interval for each detector response.

Once one has a set of optimized detection vectors for a given data set, it is then necessary to solve Eq. (29) (if we did not include S^2, then we eliminate the first column and last row from the matrix). Because our real data include noise, we also want to account for signal-to-noise of each measurement (fitting by minimizing χ^2). Then, our matrix takes on the following form:

$$\begin{pmatrix} \rho_0^{(\theta,S)} \\ \rho_1^{(\theta,S)} \\ \rho_2^{(\theta,S)} \\ \rho_3^{(\theta,S)} \\ \rho_4^{(\theta,S)} \end{pmatrix} = \begin{pmatrix} [\vec{r}_0]_1/\sigma_{R1\rho,60} & [\vec{r}_1]_1/\sigma_{R1\rho,60} & \cdots & \cdots & \cdots \\ \vdots & \ddots & \cdots & \cdots & \cdots \\ [\vec{r}_0]_4/\sigma_{R1,500} & [\vec{r}_1]_4/\sigma_{R1,500} & \ddots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ [\vec{r}_0]_7/\sigma_{S^2} & [\vec{r}_1]_7/\sigma_{S^2} & \cdots & \cdots & \ddots \end{pmatrix}^{-1} \begin{pmatrix} R_{1\rho,60}/\sigma_{R1\rho,60} \\ \vdots \\ R_{1,500}/\sigma_{R1,500} \\ \vdots \\ \left(1-S^2\right)/\sigma_{S^2} \end{pmatrix}. \tag{38}$$

$[\vec{r}_n]_i$ indicate the ith element of detection vector $\vec{r}_n$. As one can see, each relaxation rate constant and $1 - S^2$ have been divided by their standard deviation. In this example, if data were missing, the rate was set to 0 and σ was set to 10^{10} to effectively remove that data from the fit. Similarly, the corresponding row of the matrix containing the detection

vectors, $\vec{r}_n$, has also been divided by the same standard deviation (elements have been omitted to save space). Instead of finding the inverse of the matrix, we use a linear solver ("lsqlin" as implemented in MATLAB[41,44]) to additionally enforce that the resulting responses, $\rho_n^{(\theta,S)}$, are non-negative [we do allow $\rho_0^{(\theta,S)}$ to be negative since it has significant regions of negative sensitivity, see Fig. 12(a)]. This forces the fitted data within the region allowed by the detection vectors (otherwise, noise pushing experimental data outside the allowed region is sometimes fitted with negative $\rho_n^{(\theta,S)}$ so that $\rho_n^{(\theta,S)}$ are noisier).

We may also obtain confidence intervals for the responses, $\rho_n^{(\theta,S)}$, using a Monte-Carlo approach.[45] Once the initial responses are obtained, we back calculate the relaxation-rate constants. For each calculated rate constant, we then add pseudo-random noise selected from a normal distribution with the experimentally determined standard deviation for that rate constant. The modified set of rate constants can be re-fit as before, and the process is repeated (200 times in this example). Then, one may obtain a desired confidence interval, C, by sorting each set of $\rho_n^{(\theta,S)}$ and finding the $(1/2 - C/2)$ Nth and $(1/2 + C/2)$ Nth elements, where N is the number of repetitions taken, yielding the lower and upper bounds for the confidence interval (we take $C = 0.95$ in our example). Note that this approach to error analysis assumes that the error on the experimental data is normally distributed but makes no assumption on the distribution on the error of the detector responses so that error bars are not necessarily symmetric and the relationships between confidence intervals for the normal distribution will not hold (for example, the 95% confidence interval is not twice as wide as the 68% confidence interval).

The results of our analysis of ubiquitin dynamics are shown in Fig. 12, where the sensitivities of the five detectors used for dynamics analysis are shown in (a), and the residue specific responses of the detectors, $\rho_n^{(\theta,S)}$, are shown in (b).

The amplitudes of each residue-specific detector response are indicative of the total amplitude of motion in the sensitive range of that detector [see Eq. (12)]. However, this relationship does not allow exact quantification of the amplitude of motion. If motion is concentrated where the detector is most sensitive, then one obtains a higher detector response than if most motion occurs where the detector is less sensitive. However, if there is a uniform distribution of motion, then the response of each detector gives the total amplitude of motion for a range of correlation times having a width of Δz_n, the effective detector width, and the sum of all detectors (if $\rho_0^{(\theta,S)}$ is included) yields the total amplitude of motion, $1 - S^2$.

An important property of the detectors is that they allow analysis of motion without making any assumptions about the number of exponential terms in the correlation function describing that motion. This means that detector responses, $\rho_n^{(\theta,S)}$, should be consistent with any dynamics analysis that uses an explicit model of motion, in other words, an analysis with a fixed number of exponential terms, as long as that model yields a good fit of the experimental data. For example, R_1 and S^2 data were previously fit to a bi-exponential correlation function.[11] This is equivalent to having a distribution of motion that is the sum of two δ-functions so that we may calculate the detector response corresponding to this explicit model of motion as follows [see Eq. (12)]:

$$\begin{aligned}(1 - S^2)\theta(z) &= (1 - S_f^2) * \delta(z = \log_{10}(\tau_f/s)) \\ &\quad + S_f^2(1 - S_s^2) * \delta(z = \log_{10}(\tau_s/s)), \\ \rho_n^{(\theta,S)} &= (1 - S_f^2)\rho_n\left(\log_{10}(\tau_f/s)\right) \\ &\quad + S_f^2(1 - S_s^2)\rho_n\left(\log_{10}(\tau_s/s)\right).\end{aligned} \tag{39}$$

Similarly, $R_{1\rho}$ data were previously fit to a mono-exponential function[43] so that from this model, the following detector response can be calculated:

$$\rho_n^{(\theta,S)} = S_f^2(1 - S_s^2)\rho_n\left(\log_{10}(\tau_c/s)\right). \tag{40}$$

Figure 13 shows the detector analysis of experimental data and detector responses calculated from explicit models using Eqs. (39) and (40). Very good agreement is obtained, with only small deviations arising in the detector responses. In some cases, the relative error on $\rho_1^{(\theta,S)}$ and $\rho_2^{(\theta,S)}$ is notably higher than that for $(1 - S_f{}^2)$ and $(1 - S_s{}^2)$ previously reported.[11] This is in part because the direct measurement of S^2 via REDOR stabilizes the model-free fit parameters (see the work of Haller and Schanda, Fig. 8[11]), whereas in the detectors approach, the direct measurement of S^2 only contributes to the determination of $\rho_0^{(\theta,S)}$ (also, we report a 95% confidence interval versus one standard deviation, i.e., 68% confidence interval, reported by Haller and Schanda).

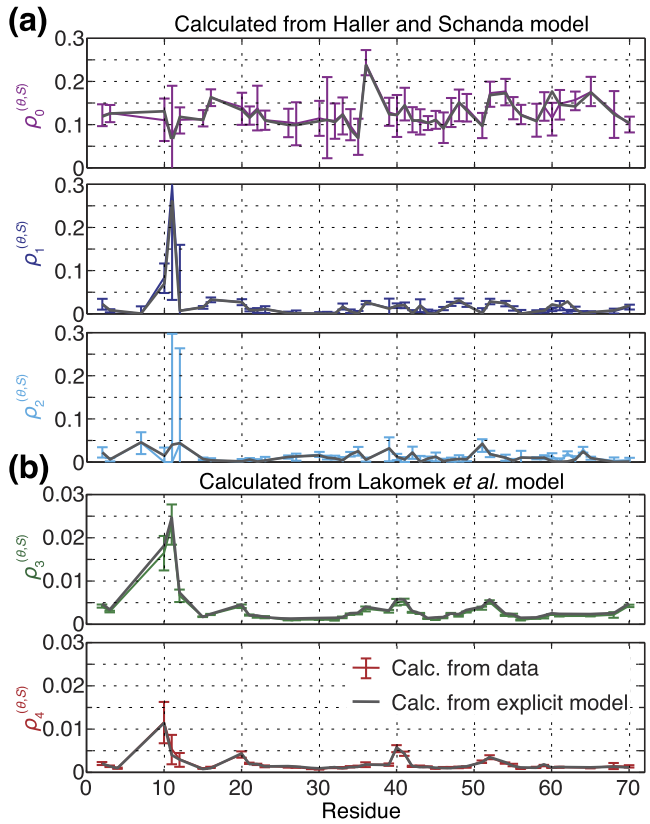

FIG. 13. Comparison of detector responses calculated directly from data (colored lines with error bars) and calculated from explicit models (black lines). (a) shows detectors calculated from a bi-exponential model from Ref. 11 using Eq. (39). (b) shows detectors calculated from a mono-exponential model from Ref. 43 using Eq. (40). Note that in (a) and (b), data are only shown for residues that had both explicit model results and detector results available.

Although a detector analysis alone will be consistent with an explicit model that leads to good fits of experimental data, one may use the detectors as a way to identify alternative interpretations of the dynamics behavior. For example, where $R_{1\rho}$ data were modeled using one correlation time,[43] we see from detector responses, $\rho_3^{(\theta,S)}$ and $\rho_4^{(\theta,S)}$, that alternative models could have a distribution of motion that covers a wide range of correlation times (including the 1-2 μs range that was identified, near where $\rho_3^{(\theta,S)}$ and $\rho_4^{(\theta,S)}$ overlap) or could even have a distribution of motion that covers parts of each of the ranges where $\rho_3^{(\theta,S)}$ and $\rho_4^{(\theta,S)}$ are sensitive, while having no motion in the 1-2 μs range. Similarly, the R_1 and S^2 data used here were previously modeled with two correlation times, one usually being shorter than 100 ps and one being longer than 10 ns. However, from the responses of $\rho_1^{(\theta,S)}$ and $\rho_2^{(\theta,S)}$, we can clearly see that another reasonable model could include significant motion in the ∼1 ns to ∼10 ns range.

If we interpret the dynamics response under the assumption that one finds some motion at all ranges of correlation times, then the detector responses are primarily due to motion where the detectors are most sensitive. With this in mind, one may identify potential artifacts of the explicit model. For example, if we consider residue 11Thr which exhibits the fastest R_1 relaxation in the protein, we notice that $(1 - S^2)$ is 0.4, compared to a median $(1 - S^2)$ of 0.15,[11] whereas R_1 at 850 MHz is 0.90 s^{-1}, compared to a median of 0.033 s^{-1}.[7] Since R_1 is proportional to $(1 - S^2)$ [Eqs. (6) and (7)], one expects these ratios to be similar, but in fact, the $(1 - S^2)$ ratio is 2.7:1 and the R_1 ratio is 27:1 [similar behavior was also observed for HET-s(218-289) at turns between β-sheets[33]]. The detector analysis indicates that a significant portion of $(1 - S^2)$ can be due to a somewhat uniform motion at short correlation times where the R_1 experiments are not sensitive, whereas the large change in the R_1 rate constant can be the result of more motion specifically where the R_1 experiments are sensitive (as exhibited by relatively uniform values of $\rho_0^{(\theta,S)}$, whereas the ratio of $\rho_1^{(\theta,S)}$ at residue 11Thr to the median value of $\rho_1^{(\theta,S)}$ is 23:1). On the other hand, the explicit model explains the inconsistency in these ratios by fitting 11Thr to a much longer correlation time (τ_f) than most other residues (760 ps versus a median of the log of τ_f of 39 ps).[11] As shown in Fig. 13, this interpretation is still consistent with the detector analysis, but in this interpretation, residues fitted to short correlation times must have virtually no motion where R_1 experiments (and $\rho_1^{(\theta,S)}$, $\rho_2^{(\theta,S)}$) are most sensitive, otherwise these motions would significantly change the R_1 rate constants. A complete lack of motion for this range in most of the protein is an unlikely situation and so is likely to be an artifact of the analysis method (as was also previously shown[14]).

V. CONCLUSIONS

In this study, we introduce a method for the visualization of the dynamic information in a set of NMR experiments and develop the detectors method for dynamics analysis: linear combinations of NMR dynamics data that can be used to characterize motion. This allows us to quantify how much motion occurs for a range of correlation times, where that range is defined by the sensitivity of each detector. We introduced a reduced space of relaxation rate constants, which may be used for selecting detection vectors that can optimally fit relaxation rate constants, while yielding well-separated and non-negative sensitivities. We also find that using different normalization schemes (Sec. III G), we can estimate the total amplitude of motion for a range of correlation times or the amplitude of the distribution function (with the accuracy depending on how motion is distributed). We have furthermore developed the DIFRATE software to aid in the application of our proposed method and demonstrated its usage on ubiquitin dynamics data. We show that previous models of the ubiquitin data are consistent with the detector analysis presented here; however, the new analysis can be interpreted much more broadly than the explicit models.

An important advantage to the methodology presented here is that it allows direct comparison to other dynamics methods, especially computational methods. Detector responses have a clear definition, given by Eq. (12), so that if one obtains a trajectory, for example, from molecular dynamics (MD) simulation, it is possible to estimate the motional distribution to determine the quality of the simulation by comparing the simulated and experimental detector responses. Furthermore, the sensitivity of the detectors clearly defines what ranges of time scales on which one obtains dynamics information. Then, one can determine in what regions of a protein and on what time scales the MD trajectory is accurate. Not only can one accept or reject a trajectory based on such an analysis but also it may be further possible to optimize force fields, using the NMR detectors as target data.[46–48]

Furthermore, one may be able to develop a model of motion from the MD trajectory, if detector responses are at least somewhat similar, and apply this model to interpret the experimental detector response. On the other hand, if an oversimplified model of the correlation function, such as the extended model-free approach, is used to fit the NMR data and to fit the MD trajectory, the biasing of the two methods may result in very different behavior in the fitted parameters, leading to disagreement even when the MD is accurate.

Although the proposed method avoids artifacts resulting from incorrect models of motion,[14] the responses of the detectors, $\rho_n^{(\theta,S)}$, do not directly give us a specific model of motion or a quantitative measurement of either correlation times or amplitudes. This is because this information is not contained in the experimental data. To obtain a quantitative description of the internal protein motion, one must take further assumptions about the type of motion (and justify those assumptions). Consider how one may do this—if the same residue-correlated trends are found in several detectors, one could assume that this describes a distribution of motion that covers the range of all of those detectors. Then, one may use the equal-integral normalization method and quantify the distribution function, $(1 - S^2)\theta(z_0)$, at the centers of each detector (z_0). At this point, one may consider possible distribution functions which are consistent with the detector responses (a number of models are described elsewhere[19]). Similarly, a strong response of one or two detectors, where the other detectors are clearly not responding, can be indicative of motion that may be described by a single correlation time, in which case it may make sense to re-introduce the simple model-free approach to obtain a

more exact correlation time and amplitude of motion. Still, one should be careful in this case and ideally have physical justification for such a model.

Note that, in this paper, we primarily discuss designing a set of detectors using a pre-determined set of experiments. However, if one is interested in characterizing a particular range of correlation times, then one should select the set of experiments in order to obtain detectors that are sensitive in the desired range. In this way, one may be more experimentally efficient. One also sees that detectors tend to fall on certain ranges of correlation times, based on what experiments are common (e.g., R_1 and $R_{1\rho}$) so that one is motivated to develop new experiments for different detector ranges; for example, high-resolution relaxometry, introduced recently in solution-state NMR,[49] could add new detectors for different ranges of correlation times.

By eliminating assumptions about the model complexity, we obtain a description of the internal protein dynamics that may be broadly interpreted. After initial results are obtained, it may be possible to make further assumptions about the model of motion. However, initially assuming a particular model without physical justification may cripple ones ability to properly interpret results. Therefore, we conclude that the approach presented here can be the basis of an unbiased dynamics analysis of NMR relaxation data.

SUPPLEMENTARY MATERIAL

See supplementary material for a glossary of terms introduced here and for tables of the results shown in Figs. 12 and 13.

ACKNOWLEDGMENTS

The authors would like to thank N. Lakomek, S. Riniker, and F. Ferrage for helpful discussions. This work was supported by the ETH Zurich and the Swiss National Science Foundation (Grant Nos. 200020_159707 and 200020_146757). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 741863, FASTER).

[1] J. R. Lewandowski, H. J. Sass, S. Grzesiek, M. Blackledge, and L. Emsley, J. Am. Chem. Soc. **133**(42), 16762 (2011).
[2] A. G. Redfield, IBM J. Res. Dev. **1**(1), 19 (1957).
[3] A. G. Redfield, Adv. Magn. Opt. Reson. **1**, 1 (1965).
[4] G. Lipari and A. Szabo, J. Am. Chem. Soc. **104**(17), 4559 (1982).
[5] G. M. Clore, A. Szabo, A. Bax, L. E. Kay, P. C. Driscoll, and A. M. Gronenborn, J. Am. Chem. Soc. **112**, 4989 (1990).
[6] V. Chevelkov, U. Fink, and B. Reif, J. Biomol. NMR **45**(1-2), 197 (2009).
[7] P. Schanda, B. H. Meier, and M. Ernst, J. Am. Chem. Soc. **132**(45), 15957 (2010).
[8] B. Halle and H. Wennerström, J. Chem. Phys. **75**(4), 1928 (1981).
[9] B. Halle, J. Chem. Phys. **131**, 224507 (2009).
[10] T. Zinkevich, V. Chevelkov, B. Reif, K. Saalwächter, and A. Krushelnitsky, J. Biomol. NMR **57**(3), 219 (2013).
[11] J. D. Haller and P. Schanda, J. Biomol. NMR **57**(3), 263 (2013).
[12] L. Mollica, M. Baias, J. R. Lewandowski, B. J. Wylie, L. J. Sperling, C. M. Rienstra, L. Emsley, and M. Blackledge, J. Phys. Chem. Lett. **3**(23), 3657 (2012).
[13] J. M. Lamley, M. J. Lougher, H. J. Sass, M. Rogowski, S. Grzesiek, and J. R. Lewandowski, Phys. Chem. Chem. Phys. **17**, 21997 (2015).
[14] A. A. Smith, M. Ernst, and B. H. Meier, Angew. Chem., Int. Ed. **129**(44), 13778 (2017).
[15] T. Smith and J. Guild, Trans. Opt. Soc. **33**(3), 73 (1931).
[16] G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae* (Wiley, 2000).
[17] G. Lipari and A. Szabo, J. Am. Chem. Soc. **104**(17), 4546 (1982).
[18] D. M. Brink and G. R. Satchler, *Angular Momentum* (Clarendon Press, Oxford, 1968).
[19] P. A. Beckmann, Phys. Rep. **171**(3), 85 (1988).
[20] P. Schanda and M. Ernst, Prog. Nucl. Magn. Reson. Spectrosc. **96**, 1 (2016).
[21] U. Haeberlen and J. S. Waugh, Phys. Rev. **185**(2), 420 (1969).
[22] R. Kurbanov, T. Zinkevich, and A. Krushelnitsky, J. Chem. Phys. **135**(18), 184104 (2011).
[23] P. Rovo and R. Linser, J. Phys. Chem. B **121**(25), 6117 (2017).
[24] J. Peng and G. Wagner, J. Magn. Reson. **98**(2), 308 (1992).
[25] M. L. Gill, R. A. Byrd, and A. G. I. Palmer, Phys. Chem. Chem. Phys. **18**(8), 5839 (2016).
[26] N. A. Farrow, O. Zhang, A. Szabo, D. A. Torchia, and L. E. Kay, J. Biomol. NMR **6**(2), 153 (1995).
[27] C. H. Papavoine, M. L. Remerowski, L. M. Horstink, R. N. Konings, C. W. Hilbers, and F. J. van de Ven, Biochemistry **36**(13), 4015 (1997).
[28] K. W. Wagner, Ann. Phys. (Ger.) **345**(5), 817 (1913).
[29] H. Akaike, IEEE Trans. Autom. Control **19**(6), 716 (1974).
[30] C. M. Hurvich and C.-L. Tsai, J. Time Ser. Anal. **14**(3), 271 (1993).
[31] J. Shang and J. Cavanaugh, Comput. Stat. Data Anal. **52**(4), 2004 (2008).
[32] I. Solomon, Phys. Rev. **99**(2), 559 (1955).
[33] A. A. Smith, E. Testori, R. Cadalbert, B. H. Meier, and M. Ernst, J. Biomol. NMR **65**(3-4), 171 (2016).
[34] J. R. Lewandowski, J. Sein, H. J. Sass, S. Grzesiek, M. Blackledge, and L. Emsley, J. Am. Chem. Soc. **132**(24), 8252 (2010).
[35] D. M. LeMaster, J. Biomol. NMR **6**(4), 366 (1995).
[36] D. M. LeMaster, J. Am. Chem. Soc. **121**(8), 1726 (1999).
[37] S. N. Khan, C. Charlier, R. Augustyniak, N. Salvi, V. Dejean, G. Bodenhausen, O. Lequin, P. Pelupessy, and F. Ferrage, Biophys. J. **109**(5), 988 (2015).
[38] P. Schanda, B. H. Meier, and M. Ernst, J. Magn. Reson. **210**(2), 246 (2011).
[39] A. Barth, Biochim. Biophys. Acta **1767**(9), 1073 (2007).
[40] L. Yao, B. Vogeli, J. Ying, and A. Bax, J. Am. Chem. Soc. **130**(49), 16518 (2008).
[41] T. Mathworks, MATLAB Release 2013b, 2013.
[42] A. A. Smith, M. Ernst, and B. H. Meier, Distortion Free Relaxation Analysis Technique software, URL: http://difrate.sourceforge.net, 2018.
[43] N. A. Lakomek, S. Penzel, A. Lends, R. Cadalbert, M. Ernst, and B. H. Meier, Chemistry **23**(39), 9425 (2017).
[44] T. Mathworks, MATLAB and Optimization Toolbox Release 2013b, 2013.
[45] N. Metropolis and S. Ulam, J. Am. Stat. Assoc. **44**(247), 335 (1949).
[46] A. D. Mackerell, Jr., J. Comput. Chem. **25**(13), 1584 (2004).
[47] A. D. Mackerell, Jr., M. Feig, and C. L. Brooks, III, J. Comput. Chem. **25**(11), 1400 (2004).
[48] W. F. van Gunsteren and H. J. C. Berendsen, Angew. Chem., Int. Ed. **29**, 992 (1990).
[49] C. Charlier, S. N. Khan, T. Marquardsen, P. Pelupessy, V. Reiss, D. Sakellariou, G. Bodenhausen, F. Engelke, and F. Ferrage, J. Am. Chem. Soc. **135**(49), 18665 (2013).
[50] T. Gullion and J. Schaefer, J. Magn. Res. **81**(1), 196 (1989).