Supplementary information for:

Reducing bias in the analysis of solution-state NMR data with dynamics detectors

Albert A. Smith[1,2], Matthias Ernst[1], Beat H. Meier[1], Fabien Ferrage[3]

[1]*ETH Zurich, Physical Chemistry, Vladimir-Prelog-Weg 2, 8093 Zurich, Switzerland*
[2]*Present address: Universität Leipzig, Insitut für Medizinische Physik und Biophysik, Härtelstraße 16-18, 04107 Leipzig, Germany*
[3]*Laboratoire des biomolécules, LBM, Département de chimie, École normale supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France.*

M.E. : maer@ethz.ch
B.M. : beme@ethz.ch
F.F. : Fabien.Ferrage@ens.fr

Table of Contents

Glossary of terms

Name	Symbol	Units	Description
Correlation time	τ_c	s	Correlation time of some motion in the system.
Log-correlation time	z	unitless (vs. 1 s)	Base-10 logarithm of the correlation time, given by $\log_{10}(\tau_c / 1\,\mathrm{s})$.
Rotational diffusion correlation time	τ_r	s	Correlation time of isotropic rotational of a molecule in solution (tumbling).
Log-τ_r	z_r	unitless (vs. 1 s)	Base-10 logarithm of the rotational correlation time, given by $z_r = \log_{10}(\tau_r / 1\,\mathrm{s})$
Effective correlation time	τ_{eff}	s	Effective correlation of an internal motion, where the molecule is undergoing tumbling with correlation time, τ_r. Given by $\tau_{eff} = \tau_r\tau_c \big/ \left(\tau_r + \tau_c\right)$.
Log-effective correlation time	z_{eff}	unitless (vs. 1 s)	Base-10 logarithm of the effective correlation time, given by $\log_{10}(\tau_{eff} / 1\,\mathrm{s})$.
Distribution of motion	$(1-S^2)\theta(z)$	unitless	Describes how motion is distributed as a function of correlation time, where z = $\log_{10}(\tau_c$ /1 s). $(1–S^2)$ gives the total amplitude of motion, so that $\theta(z)$ always integrates to one.
Distribution of internal motion	$(1-S^2)\theta(z_i)$	unitless	This is the same as the distribution of motion for solid-state analysis. In solution-state analysis, this distribution only accounts for internal motion of the molecule- in other words, tumbling of the molecule is factored out, and the log-correlation times are *not* effective correlation times (see SI section 1 for comparison of distributions).
Distribution of total motion	$\theta_{tot.}(z)$	unitless	This is the distribution of all motions for a molecule tumbling in solution, including the tumbling itself. Motion resulting from internal motion is modified to have an effective correlation time, z_{eff}, which results from the internal correlation time and the tumbling correlation time see SI section 1 for comparison of distributions).

Relaxation rate constant	$R_{\zeta}^{(\theta,S)}$	s^{-1}	The relaxation-rate constant obtained under experimental conditions denoted by ζ, for a distribution of motion $(1-S^2)\theta(z)$. May be obtained by integrating the product of the sensitivity of that rate constant, $R_{\zeta}(z)$, times the distribution of motion, $(1-S^2)\theta(z)$.
Sensitivity	$R_{\zeta}(z)$	s^{-1}	The relaxation rate constant obtained under experimental conditions denoted by ζ, for a mono-exponential correlation function, having correlation time $\tau_c = 10^z$ s, and amplitude $1-S^2 = 1$.
Solution-state sensitivity	$R_{\zeta}^{solu}(z_i)$	s^{-1}	Sensitivity of an experiment to the internal motion of a molecule, with $\tau_i = 10^{z_i} \cdot 1\,\mathrm{s}$, when the molecule is tumbling in solution. This function has one term to account for attenuation of relaxation due to rotational diffusion, and a second term to account for relaxation induced by the internal motion, given as $R_{\zeta}^{solu.}(z_i) = R_{\zeta}(z_{eff}(z_i)) - R_{\zeta}(z_r)$.
Detector	–	–	A mathematical tool used to quantify the amount of motion for a range of correlation times.
Detector sensitivity	$\rho_n(z)$	unitless	Defines how a detector responds to a particular correlation time, $\tau_c = 10^z$ s. Its value as a function of z is obtained by taking a linear combination of rate constant sensitivities (using the same linear combination as is used to obtain the detector responses).
Detector response	$\rho_n^{(\theta,S)}$	unitless	A quantity, describing the amount of motion for a particular range of correlation times, rigorously defined as the integral of the product of the detector sensitivity, $\rho_n(z)$, and the distribution of motion, $(1-S^2)\theta(z)$. Obtained by taking an appropriate linear combination of experimental rate constants (strictly speaking, by fitting a vector of the rate constants to the detection vectors, $\vec{r}_n$).
Normalized rate constant	$\Re_{\zeta}^{(\theta,S)}$	unitless	The relaxation rate constant divided by some normalization constant, c_ζ, to yield a dimensionless relaxation rate constant. For solution state relaxation, we first subtract away the relaxation rate constant obtained for an internally rigid motion, R_{ζ}^0, such that $\Re_{\zeta}^{(\theta,S)} = (R_{\zeta}^{(\theta,S)} - R_{\zeta}^0)/c_{\zeta}$.
Allowed region	–	–	For a given set of experiments, the allowed region is all sets of rate constants ($R_{\zeta}^{(\theta,S)}$) that can be obtained for any arbitrary distribution of motion, given by $(1-S^2)\theta(z)$. Usually this space is represented in terms of the $\Re_{\zeta}^{(\theta,S)}$.

Detection vector	$\vec{r}_n$	s^{-1}	A vector containing carefully chosen values of the $R_\zeta^{(\theta,S)}$, so that a vector containing the full set of experimentally determined relaxation rate constants is assumed to be a linear combination of all detection vectors, given by $\rho_1^{(\theta,S)}\vec{r}_1 + \rho_2^{(\theta,S)}\vec{r}_2 + \ldots$.
Sum of normalized rate constants	$\Sigma_\zeta \Re_\zeta^{(\theta,S)}$	unitless	Sum of all normalized rate constants for an experimental data set, used for calculating the ratio of rates. Note that for the reduced space for internal motion (solution-state), this term is replaced, often by $-\Re_\zeta^{(\theta,S)}$, where the corresponding sensitivity, $R_\zeta(z)$, remains negative for all correlation times (see main text, Eq. (22)).
Ratio of rates	κ_ζ	unitless	For experimental conditions denoted by ζ, this is the ratio of the normalized rate constants, $\Re_\zeta^{(\theta,S)}$, divided by the sum of normalized rate constants, $\Sigma\Re^{(\theta,S)}$, which is used for defining positions in the reduced space.
Reduced space	–	–	For a set of experiments, the reduced space is defined by the ratios of rates, κ_ζ, for that set of experiments. The dimensionality of this space is one less than the number of experiments- achieved by omitting one of the experiments when calculating the κ_ζ.
Reduced vector	$\vec{\kappa}$	unitless	Vector of ratios of rates, κ_ζ, defining a position in the reduced space. These positions can be used to define detection vectors, although note that the reduced vector only defines the direction of the detection vector, but not the length.
Effective width	Δz_n	unitless (vs. 1 s)	The effective width of a detector is defined as the detector integral divided by its maximum, given on a base-10 log scale. $\Delta z = \int \rho_n(z)dz \Big/ \max(\rho_n(z))$
Detector center	z_n^0	unitless (vs. 1 s)	This gives the center of the detector sensitivity, on a logarithmic scale (unitless, with reference to 1 s using a base-10 log). Defined as follows: $z_n^0 = \int z\rho_n(z)dz \Big/ \int \rho_n(z)dz$

1. Distribution of the total motion vs. distribution of internal motion

In the case of a molecule tumbling isotropically in solution, we may assume that the total correlation is a product of the correlation function of the internal motion and the correlation function of the tumbling, such that

$$\begin{aligned} C(t) &= C_{\mathrm{O}}(t)C_{\mathrm{I}}(t) \\ C_{\mathrm{O}}(t) &= \frac{1}{5}\exp(-t/\tau_{\mathrm{r}}) \end{aligned} \quad . \tag{S1}$$

Here, we write correlation function of the internal motion as

$$C_{\mathrm{I}}(t) = \frac{1}{5}\left[S^2 + (1-S^2)\int_{-\infty}^{\infty}\theta(z_{\mathrm{i}})\exp(-t/(10^{z_{\mathrm{i}}}\cdot 1\,\mathrm{s}))dz\right], \tag{S2}$$

where $(1-S^2)\theta(z_{\mathrm{i}})$ describes the distribution of internal motion (we use z_{i} to distinguish the correlation time of the internal motion from z_{eff}, which appears in the next equation as the effective correlation time). The product of the two correlation functions then yields

$$\begin{aligned} C(t) &= \frac{1}{5}\left[S^2\exp(-t/\tau_{\mathrm{r}}) + (1-S^2)\int_{-\infty}^{\infty}\theta(z_{\mathrm{i}})\exp(-t/(10^{z_{\mathrm{eff}}(z_{\mathrm{i}})}\cdot 1\,\mathrm{s}))dz_{\mathrm{i}}\right] \\ \tau_{\mathrm{eff}}(z_{\mathrm{i}}) &= \frac{(10^{z_{\mathrm{i}}}\cdot 1\,\mathrm{s})\tau_{\mathrm{r}}}{(10^{z_{\mathrm{i}}}\cdot 1\,\mathrm{s})+\tau_{\mathrm{r}}},\ \ z_{\mathrm{eff}}(z_{\mathrm{i}}) = \log_{10}(\tau_{\mathrm{eff}}(z_{\mathrm{i}})/1\,\mathrm{s}) \end{aligned} \quad . \tag{S3}$$

One notes, however, that this correlation function is still a sum of decaying exponential terms, although with modified correlation times (given by $z_{\mathrm{eff}}(z) = \log_{10}(\tau_{\mathrm{eff}}/1\,\mathrm{s})$). Therefore, we can analyze relaxation arising from such a correlation function using the detector analysis as derived for solid-state NMR,[1] however, we will not characterize the distribution of internal motion, $(1-S^2)\theta(z_{\mathrm{i}})$, but rather some distribution of the total motion, $\theta_{\mathrm{tot.}}(z)$,such that

$$C(t) = \frac{1}{5}\int_{-\infty}^{\infty}\theta_{\mathrm{tot.}}(z)\exp(-t/(10^{z}\cdot 1\,\mathrm{s}))dz \quad , \tag{S4}$$

where the correlation functions in Eqs. (S3) and (S4) are equal. Note that due to the tumbling, the total motion is isotropic, so that $(1-S^2)$ in this case equals 1, and is therefore omitted from Eq. (S4). We will see during the following derivation, that the z appearing in this equation may denote a log-effective correlation time, z_{eff}, or log-correlation time for the tumbling, z_{r}, so that we simply denote this variable as z.

Then, we would like to know the relationship between $(1-S^2)\theta(z_i)$ and $\theta_{tot.}(z)$ for a molecule tumbling with correlation time τ_r. We do so by rearrangement of Eqs. (S3) and (S4). To begin, we define $\theta'_{tot.}(z)+S^2\delta(z=z_r)=\theta_{tot.}(z)$, where $z_r=\log_{10}(\tau_r/1\,\mathrm{s})$. Inserting into Eq. (S4), we obtain

$$\begin{aligned} C(t) &= \frac{1}{5}\int_{-\infty}^{\infty}\left(\theta'_{tot.}(z)+S^2\delta(z-z_r)\right)\exp(-t/(10^z\cdot 1\,\mathrm{s}))dz \\ &= \frac{1}{5}\left[S^2\exp(-t/\tau_r)+\int_{-\infty}^{\infty}\theta'_{tot.}(z)\exp(-t/(10^z\cdot 1\,\mathrm{s}))dz\right], \end{aligned} \tag{S5}$$

Addition of the δ-function has produced the first term in Eq. (S3), so that by setting Eqs. (S3) and (S5) equal, we may obtain

$$(1-S^2)\int_{-\infty}^{\infty}\theta(z_i)\exp(-t/(10^{z_{eff}(z_i)}\cdot 1\,\mathrm{s}))dz_i=\int_{-\infty}^{\infty}\theta'_{tot.}(z)\exp(-t/(10^z\cdot 1\,\mathrm{s}))dz \tag{S6}$$

.

Then, we see that the z in the right side of this equation must be equal to $z_{eff}(z_i)$ if the two integrals are equal. Thus, we simply replace all z with z_{eff} on the right side

$$(1-S^2)\int_{-\infty}^{\infty}\theta(z_i)\exp(-t/(10^{z_{eff}(z_i)}\cdot 1\,\mathrm{s}))dz_i=\int_{-\infty}^{\infty}\theta'_{tot.}(z_{eff})\exp(-t/(10^{z_{eff}}\cdot 1\,\mathrm{s}))dz_{eff} \tag{S7}$$

,

followed by changing the integration variable on the left side to z_{eff}. To do so, we need to obtain z_i and dz_i in terms of z_{eff}, and furthermore adjust the integration bounds.

From the definition of τ_{eff} (Eq. (S3)), we start with

$$\begin{aligned} \tau_{eff} &= 10^{z_{eff}}=\frac{10^{z_i}\cdot\tau_r}{10^{z_i}+\tau_r}=\frac{10^{z_i+z_r}}{10^{z_i}+10^{z_r}} \\ 10^{z_{eff}}(10^{z_i}+10^{z_r}) &= 10^{z_i}\cdot 10^{z_r} \\ 10^{z_i} &= \frac{10^{z_{eff}}\cdot 10^{z_r}}{10^{z_r}-10^{z_{eff}}} \\ z_i &= z_{eff}+z_r-\log_{10}(10^{z_r}-10^{z_{eff}}) \\ dz_i &= dz_{eff}+dz_{eff}\frac{10^{z_{eff}}}{10^{z_r}-10^{z_{eff}}}=dz_{eff}\frac{10^{z_r}}{10^{z_r}-10^{z_{eff}}} \end{aligned} \tag{S8}$$

,

Next, we find the upper and lower bounds of the integral

Lower bound:	Upper bound:
$z_i = -\infty,$	$z_i = \infty$
$10^{z_{eff}} = \frac{10^{-\infty+z_r}}{10^{-\infty}+10^{z_r}} = 0$,	$10^{z_{eff}} = \frac{10^{\infty+z_r}}{10^{\infty}+10^{z_r}} = 10^{z_r}$,
$z_{eff} = -\infty$	$z_{eff} = z_r$

(S9)

Plugging in, we obtain

$$(1-S^2)\int_{-\infty}^{z_r}\theta(z_{eff}+z_r-\log_{10}(10^{z_r}-10^{z_{eff}}))\exp(-t/(10^{z_{eff}}\cdot 1\,\mathrm{s})\frac{10^{z_r}}{10^{z_r}-10^{z_{eff}}}dz_{eff}$$
$$=\int_{-\infty}^{\infty}\theta'_{tot.}(z_{eff})\exp(-t/(10^{z_{eff}}\cdot 1\,\mathrm{s}))dz_{eff} \quad , \tag{S10}$$

We see that we may satisfy the equality with the following definition for $\theta'_{tot.}(z)$:

$$\theta'_{tot.}(z_{eff})=\begin{cases}(1-S^2)\theta(z_{eff}+z_r-\log_{10}(10^{z_r}-10^{z_{eff}}))\frac{10^{z_r}}{10^{z_r}-10^{z_{eff}}}, & z_{eff}<z_r\\ 0 & z\geq z_r\end{cases} , \tag{S11}$$

Finally, we may calculate $\theta_{tot.}(z)$:

$$\theta_{tot.}(z)=\begin{cases}(1-S^2)\theta(z+z_r-\log_{10}(10^{z_r}-10^{z}))\frac{10^{z_r}}{10^{z_r}-10^{z}}, & z<z_r\\ S^2\delta(z-z_r) & z=z_r\\ 0 & z>z_r\end{cases} , \tag{S12}$$

Note that for the case z=z_r, this is no longer an effective correlation time, but simply the log-correlation time of the tumbling, so that we define this function in terms of a general log-correlation time, z, as opposed to z_{eff}.

One sees that the result is reasonable. A δ-function introduces the relaxation due to tumbling into the total distribution, so the integral of this term results in the correct amplitude, S^2. The effective correlation time cannot exceed the correlation time of the tumbling, so the total distribution becomes zero for $z>z_r$. At very short correlation times, the total distribution becomes equal to the distribution of internal motion ($\log_{10}(10^{z_r}-10^{z})=z_r$, $10^{z_r}/(10^{z_r}-10^{z})=1$). As the correlation time of the total distribution approaches the rotational correlation time, one uses the effective correlation time in the distribution of internal motion, and further scales up the distribution, since one integrates over a narrower range of correlation times. Fig. S1 illustrates this for two distributions:

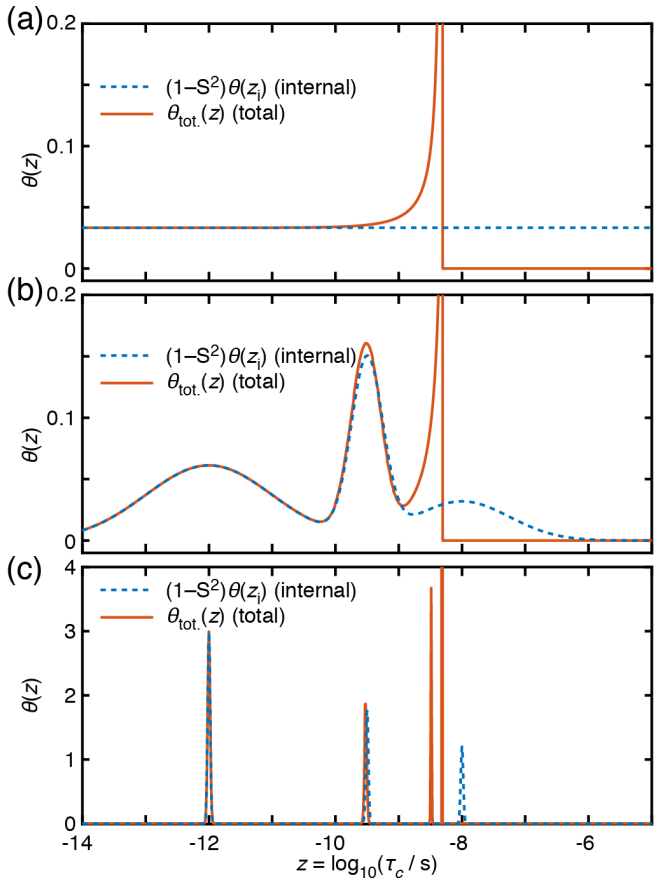

Fig. S1. Distributions of internal motion vs. distribution of total motion. Subplots (a)–(c) each show a distribution of the internal motion ($(1–S^2)\theta(z_i)$: blue, dashed line) and the resulting total distribution of motion ($\theta_{tot.}(z)$: red, solid line), assuming a rotational correlation time of $\tau_r = 4.84$ ns. (a) shows a uniform distribution for the internal motion, (b) a distribution resulting from three log-Gaussian distributions for the internal motion, and (c) shows three narrow distributions for the internal motion. Note that at the rotational correlation time, the distribution of total motion diverges to infinity (δ-function), and then falls to zero for all $z>z_r$.

2. Singular-value decomposition approach to detector optimization

2.1. *Designing the detectors*

For large data sets, the 'spaces' method of detector optimization recently developed becomes increasingly challenging, although is nonetheless very powerful for visualization of the information content of relaxation data.[1] Therefore, we introduce an alternative approach here, which utilizes reduced singular-value decomposition (SVD).[2] We begin with a matrix, **M**, which contains the normalized rate constants for a range of correlation times, for example

$$\mathbf{M} = \begin{bmatrix} \Re_\zeta(z_1) & \Re_\zeta(z_2) & \cdots & \Re_\zeta(z_n) \\ \Re_\psi(z_1) & \Re_\psi(z_2) & \cdots & \Re_\psi(z_\nu) \\ \vdots & \vdots & \ddots & \vdots \\ \Re_\xi(z_1) & \Re_\xi(z_1) & \cdots & \Re_\xi(z_1) \end{bmatrix}. \tag{S13}$$

where the ζ, φ, ξ are different experimental conditions, and the z_i are elements of a vector of correlation times (log-spaced over the full range of experiment sensitivity, see Fig. S2(a) for an example). Note that ideally, normalization is done with the standard deviation of that experiment, and in the case of multiple residues, we use the median of the standard deviation (otherwise, we normalize the sensitivity with the maximum of its absolute value).

$$\begin{aligned} \Re_\zeta(z) &= R_\zeta(z) / c_\zeta \\ c_\zeta &= \text{median}(\sigma(R_\zeta)) \end{aligned}. \tag{S14}$$

Then, SVD returns three matrices, such that

$$\mathbf{M} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}', \tag{S15}$$

where, if **M** is an *m*x*n* matrix, then **U** is an *m*x*m* unitary matrix ($\mathbf{U}^{-1}=\mathbf{U'}$, columns of **U** form an orthonormal basis), **V** is an *n*x*n* unitary matrix, and **Σ** is a diagonal *m*x*n* matrix with non-negative, real numbers on the diagonal. Here, we will typically use the truncated SVD, such that

$$\tilde{\mathbf{M}} = \mathbf{U}_t \cdot \mathbf{\Sigma}_t \cdot \mathbf{V}_t', \tag{S16}$$

where $\tilde{\mathbf{M}}$ is the closest approximation to **M,** possible with a matrix of rank *t* ($\mathbf{\Sigma}_t$ contains the *t* largest eigenvalues of **Σ**). Then, $\mathbf{U}_t$ is an *m*x*t* matrix, $\mathbf{\Sigma}_t$ is a *t*x*t* diagonal matrix, and $\mathbf{V}_t'$ is a *t*x*n* matrix.

In principle, we could define the columns of ($\mathbf{U}_t\mathbf{\Sigma}_t$) as our detection vectors (after re-normalization by the c_ζ), and the rows of $\mathbf{V}_t'$ as the corresponding sensitivities. However, we see in Fig. S2(b), that the rows of $\mathbf{V}_t'$ are not well-separated sensitivities. This is straightforward to remedy– we simply take linear combinations of the rows of $\mathbf{V}_t'$ that are optimally separated. An example of such linear combinations is shown in Fig. S2(c).

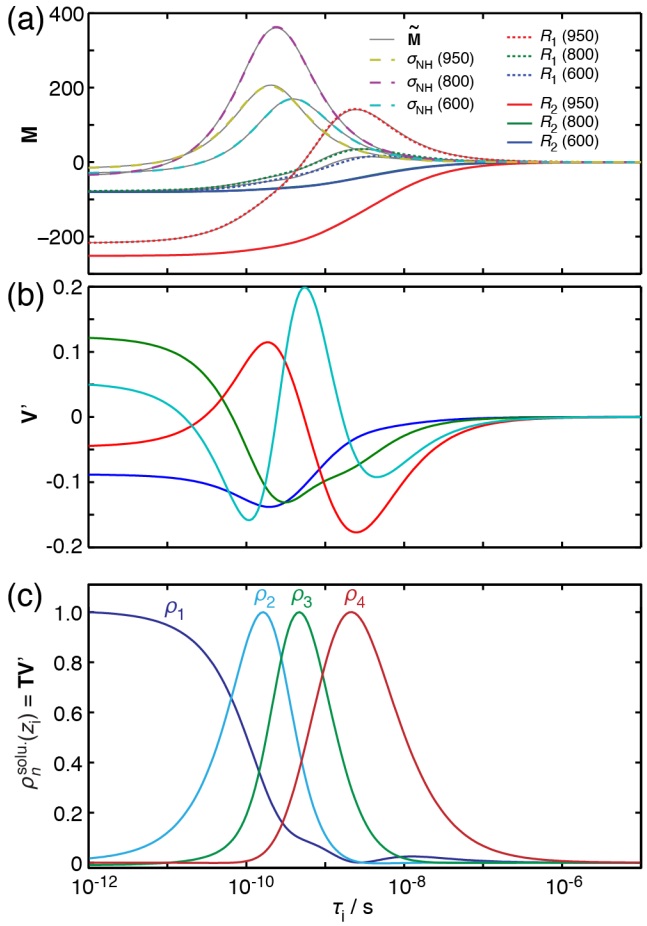

Fig. S2. Steps in the singular-value decomposition procedure. Example here are R_2, R_1, and *NOE* experiments at 600, 800, and 950 MHz. (a) shows rows of the matrix, **M**, which are the sensitivities of the different experiments for a range of correlation times, here normalized by the median standard deviation of that experiment type (taken from [3]). Note that $\tilde{\mathbf{M}}$ is also shown as grey lines (Eq. (S16), for truncated SVD of rank 4), although strongly overlaps with **M** so it is not always visible. (b) shows the rows of **V**', for a truncated SVD of rank 4. (c) shows the detector sensitivities, obtained from linear combinations of the rows of **V**', given by **TV**'.

We denote this transformation as

$$\rho_n(z_m) = \sum_{i=1}^{t} [\mathbf{T}]_{n,i} [\mathbf{V}]_{i,m} , \tag{S17}$$

where **T** is a transformation matrix, for which each row defines a linear combination of the rows in $\mathbf{V}_t'$ to yield one of the detector sensitivities, $\rho_n(z)$. Assuming that the t rows of **T** are linearly independent, then $\mathbf{T}^{-1}$ is well-defined, so that one obtains

$$\tilde{\mathbf{M}} = \mathbf{U}_t \cdot \mathbf{\Sigma}_t \cdot \mathbf{T}^{-1} \cdot \mathbf{T} \cdot \mathbf{V}_t' . \tag{S18}$$

If we renormalize $\tilde{\mathbf{M}}$, by multiplying by a diagonal matrix, **c**, which has the normalization constants, c_ζ, along its diagonal we can obtain a matrix that contains the detection vectors along its columns, here referred to as **r**.

$$\mathbf{r} = \mathbf{c} \cdot \mathbf{U}_t \cdot \mathbf{\Sigma}_t \cdot \mathbf{T}^{-1} . \tag{S19}$$

Then, the vector of experimental rate constants is fitted to

$$\min \sum_{\zeta} \sum_{n} \frac{\left(R_{\zeta}^{exp.} - [\mathbf{r}]_{\zeta,n} \rho_n^{(\theta,S)} \right)^2}{\sigma(R_{\zeta})^2} . \tag{S20}$$

where the $\rho_n^{(\theta,S)}$ are variable, or, in matrix form, we solve

$$\min \left| \begin{pmatrix} [\mathbf{r}]_{\zeta,1}/\sigma(R_{\zeta}) & [\mathbf{r}]_{\zeta,2}/\sigma(R_{\zeta}) & \cdots & [\mathbf{r}]_{\zeta,n}/\sigma(R_{\zeta}) \\ [\mathbf{r}]_{\psi,1}/\sigma(R_{\psi}) & [\mathbf{r}]_{\psi,2}/\sigma(R_{\psi}) & \cdots & [\mathbf{r}]_{\psi,n}/\sigma(R_{\psi}) \\ \vdots & \vdots & \ddots & \vdots \\ [\mathbf{r}]_{\xi,1}/\sigma(R_{\varphi}) & [\mathbf{r}]_{\xi,2}/\sigma(R_{\varphi}) & \cdots & [\mathbf{r}]_{\xi,n}/\sigma(R_{\varphi}) \end{pmatrix} \cdot \begin{pmatrix} \rho_1^{(\theta,S)} \\ \rho_2^{(\theta,S)} \\ \vdots \\ \rho_n^{(\theta,S)} \end{pmatrix} - \begin{pmatrix} R_{\zeta}^{exp.}/\sigma(R_{\zeta}) \\ R_{\psi}^{exp.}/\sigma(R_{\psi}) \\ \vdots \\ R_{\xi}^{exp.}/\sigma(R_{\xi}) \end{pmatrix} \right|^2 \tag{S21}$$

where the $|\ldots|^2$ indicates the 2-norm. Note that we restrict the $\rho_n^{(\theta,S)}$ such that $\min \rho_n(\boldsymbol{z}) \le \rho_n^{(\theta,S)} \le \max \rho_n(\boldsymbol{z})$ when solving.

We still must optimize **T**, to give well-separated detector sensitivities. We do so by choosing a target function for each detector ($\rho_n^{target}(\boldsymbol{z}_m)$), and minimizing

$$\sum_{m} \left| \left(\sum_{i=1}^{t} [\mathbf{T}]_{n,i} [\mathbf{V}]_{i,m} \right) - \rho_n^{target}(\boldsymbol{z}_m) \right|^2 , \tag{S22}$$

This has been implemented in the DIFRATE software,[4] as an interactive program with several options for the target function ('SVD_inter.m'), or as a command-line function that takes any user-defined target function ('SVD_target.m').

2.2. *Standard deviation of detectors determined from the singular values*

We can estimate the standard deviation of each detector for a given data set, using the singular values. One notes that, if we neglect the requirement that $\min \rho_n(\boldsymbol{z}) \le \rho_n^{(\theta,S)} \le \max \rho_n(\boldsymbol{z})$ then the solution to Eq. (S21) is given by

$$\begin{pmatrix} \rho_1^{(\theta,S)} \\ \rho_2^{(\theta,S)} \\ \vdots \\ \rho_n^{(\theta,S)} \end{pmatrix} = \begin{pmatrix} [\mathbf{r}]_{\zeta,1}/\sigma(R_{\zeta}) & [\mathbf{r}]_{\zeta,2}/\sigma(R_{\zeta}) & \cdots & [\mathbf{r}]_{\zeta,n}/\sigma(R_{\zeta}) \\ [\mathbf{r}]_{\psi,1}/\sigma(R_{\psi}) & [\mathbf{r}]_{\psi,2}/\sigma(R_{\psi}) & \cdots & [\mathbf{r}]_{\psi,n}/\sigma(R_{\psi}) \\ \vdots & \vdots & \ddots & \vdots \\ [\mathbf{r}]_{\varphi,1}/\sigma(R_{\varphi}) & [\mathbf{r}]_{\varphi,2}/\sigma(R_{\varphi}) & \cdots & [\mathbf{r}]_{\varphi,n}/\sigma(R_{\varphi}) \end{pmatrix}^{-1} \cdot \begin{pmatrix} R_{\zeta}^{exp.}/\sigma(R_{\zeta}) \\ R_{\psi}^{exp.}/\sigma(R_{\psi}) \\ \vdots \\ R_{\varphi}^{exp.}/\sigma(R_{\varphi}) \end{pmatrix} . \tag{S23}$$

Since this results in a simple linear combination of the experimental rate constants, we can use the usual propagation-of-error rules to obtain the standard deviation of the detectors (if

$r = ax + by + cz$, then $\sigma^2(r) = a^2\sigma^2(x) + b^2\sigma^2(y) + c^2\sigma^2(z)$ assuming zero covariance). If we take **M** to be the inverse matrix in SI Eq. (S23), then variances for each detector are given by

$$\sigma^2(\rho_n) = \sum_{\zeta} (\mathbf{M}_{n,\zeta})^2 (1)^2 . \tag{S24}$$

The experiments are already normalized by their own standard deviations, so that the variance contribution from each experiment to the detector is just $(1)^2$. Then the variance for each detector is simply the sum of the elements in the corresponding row of the inverse matrix (sum over all experiments, ζ). If we substitute the c_ζ for the $\sigma(R_\zeta)$ in this matrix (the c_ζ are just the median of the residue specific $\sigma(R_\zeta)$, so this will change the result slightly, but is a good way to understand the general behavior- see SI Eq. (S19)), the matrix inverse is given by

$$\mathbf{M} = \left(\mathbf{U}_t \cdot \mathbf{\Sigma}_t \cdot \mathbf{T}^{-1}\right)^{-1} = \mathbf{T} \cdot \mathbf{\Sigma}_t^{-1} \mathbf{U}_t' . \tag{S25}$$

Then, the variance for a given detector is given by the 2-norm of the corresponding row of this matrix:

$$\sigma^2(\rho_n^{(\theta,S)}) = \sum_{\zeta} \left[\mathbf{T}\Sigma_t^{-1}\mathbf{U}_t' \right]_{n,\zeta}^2 \tag{S26}$$

We may simplify this equation by first separating the matrix product into two parts ($\mathbf{T}\Sigma_t^{-1}$ and $\mathbf{U}_t'$), inserting a sum over the t singular values, and multiplying out $\mathbf{T}\Sigma_t^{-1}$ (which is straightforward since Σ_t^{-1} is diagonal)

$$\begin{aligned} \sigma^2(\rho_n^{(\theta,S)}) &= \sum_{\zeta} \left[\sum_{i=1}^{t} [\mathbf{T}\Sigma_t^{-1}]_{n,i} [\mathbf{U}_t']_{i,\zeta} \right]^2 \\ \left[\mathbf{T}\Sigma_t^{-1} \right]_{n,i} &= \mathbf{T}_{n,i} [\Sigma_t^{-1}]_{i,i} \\ \sigma^2(\rho_n^{(\theta,S)}) &= \sum_{\zeta} \left[\sum_{i=1}^{t} \mathbf{T}_{n,i} [\Sigma_t^{-1}]_{i,i} [\mathbf{U}_t']_{i,\zeta} \right]^2 \end{aligned} \tag{S27}$$

We then expand the squared term, to obtain

$$\begin{aligned} \sigma^2(\rho_n^{(\theta,S)}) &= \sum_{\zeta} \left[\sum_{i=1}^{t} \mathbf{T}_{n,i} [\Sigma_t^{-1}]_{i,i} [\mathbf{U}_t']_{i,\zeta} \right] \left[\sum_{j=1}^{t} \mathbf{T}_{n,j} [\Sigma_t^{-1}]_{j,j} [\mathbf{U}_t']_{j,\zeta} \right] \\ &= \sum_{i=1}^{t} \sum_{j=1}^{t} \mathbf{T}_{n,i} \mathbf{T}_{n,j} [\Sigma_t^{-1}]_{i,i} [\Sigma_t^{-1}]_{j,j} \sum_{\zeta} [\mathbf{U}_t]_{\zeta,i} [\mathbf{U}_t]_{\zeta,j} \end{aligned} \tag{S28}$$

The rearrangement of the summation order allows us to first sum over the ζ, and because the columns of $\mathbf{U}_t$ are orthonormal, this yields 1 for the inner sum, $\sum_\zeta [\mathbf{U}_t]_{\zeta,i}[\mathbf{U}_t]_{\zeta,j}$, if $i=j$, and 0 otherwise. Therefore, we obtain for the variance

$$\sigma^2(\rho_n^{(\theta,S)}) = \sum_{i=1}^{t} \left(\mathbf{T}_{n,i}[\Sigma_t^{-1}]_{i,i}\right)^2 \tag{S29}$$

Then, the variance of each detector depends on the squared inverse of the singular values with weighting determined by the corresponding row of the **T** matrix. Note that this slightly over-estimates the error, because when actually fitting, one enforces that $\min \rho_n(\boldsymbol{z}) \le \rho_n^{(\theta,S)} \le \max \rho_n(\boldsymbol{z})$. Without this requirement, experimental noise can push the detector responses outside this range, so that enforcing this requirement removes any such noise that would push the detector responses outside this range.

2.3. *Selecting the number of detectors*

Selecting greater or fewer numbers of detectors has a number of effects. More detectors will yield a better fit of the initial data set. It will also allow one to obtain detector sensitivities covering a narrower range of correlation times. However, inclusion of more detectors also means that one will have smaller singular values in the matrix Σ_t, which we can see in Eq. (S29), yields higher error for the detector responses because of the inclusion of inverse of the singular values, $[\Sigma]_{i,i}^{-1}$. For an example, we take R_1 and NOE rate constants at 600, 800, and 950 MHz, and R_2 rate constants at 950 MHz, assuming a rotational correlation time of τ_r=4.84 ns. We then calculate, for different numbers of detectors, the quality of fit of each rate constant vs. correlation time, an optimized set of detectors, and the resulting standard deviation of each detector (we will assume that the standard deviation of each measurement 5% of the maximum of the absolute value of the sensitivity). One sees that, in this case, the fit converges when using ~4 detectors, whereas using more detectors yields negligible improvement in the fit, and the standard deviation for each detector grows significantly.

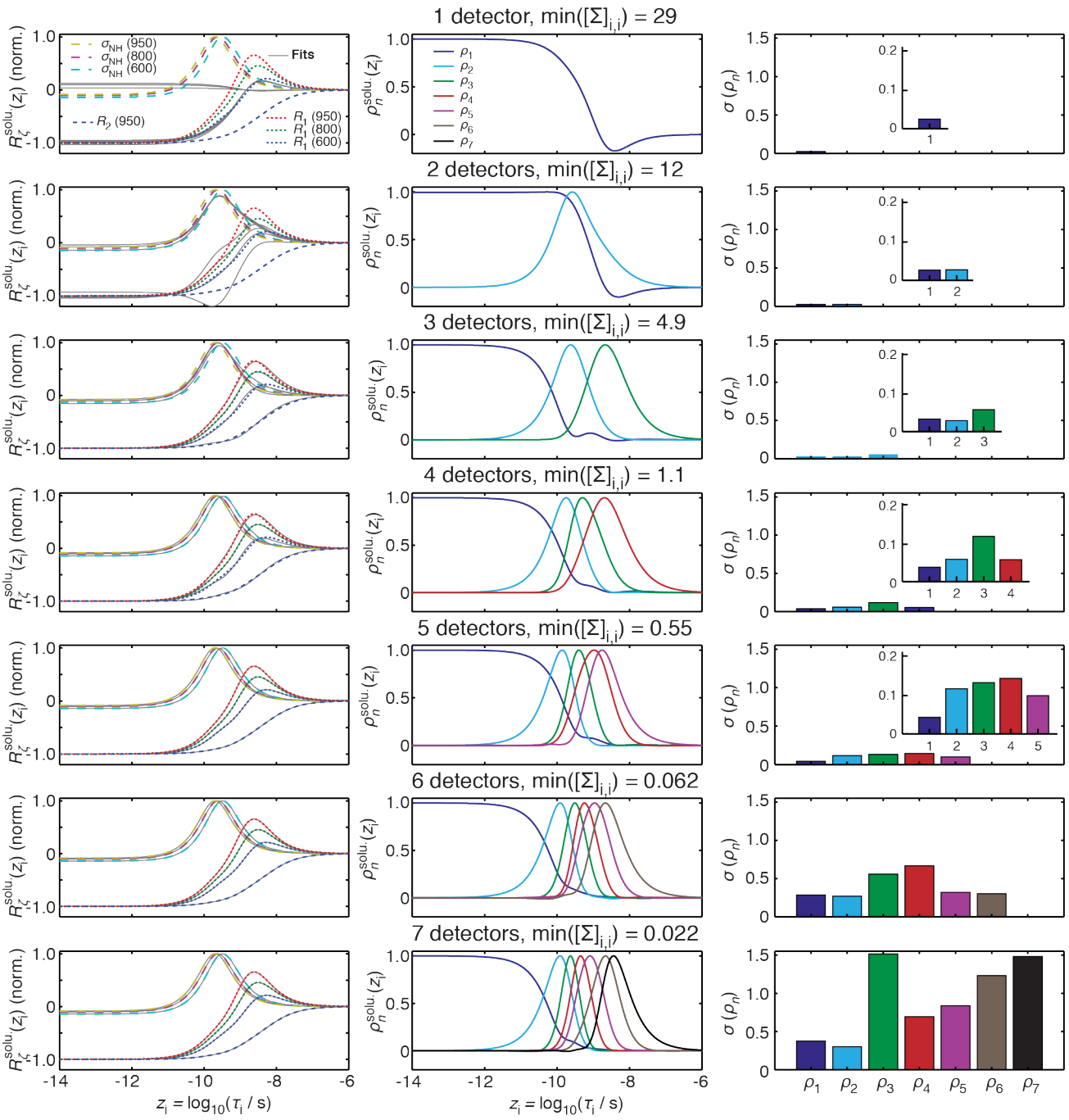

Fig. S3. Selection of the number of detectors. Each row uses the number of detectors indicated above the row to fit the set of experiment rate constants. The left column shows the sensitivities of 3 R_1 rate constants (600, 800, 950 MHz), 3 NOE rate constants (600, 800, 950 MHz), and 1 R_2 rate constant (950 MHz) as colored, dashed lines. Fits of those rate constants using the indicated number of detectors, is shown as solid, grey lines. The middle column shows an optimized set of detectors. The right column shows the standard deviation of each detector, assuming a standard deviation for each rate constant that is 5% of the maximum of the absolute value of the rate constant sensitivity. Inset on some plots shows the same information, scaled up for visibility.

3. Ubiquitin analysis at 2 fields

We calculated Ubiquitin dynamics analysis for a data at two fields (600 and 800 MHz, with R_1, R_2, and NOE data). This is analyzed with 4 detectors, assuming an overall rotational correlation time, τ_r, of 4.84 ns. The results are shown in Fig. S4. Note that we do not treat exchange in this example (compare to main text Fig. 9)

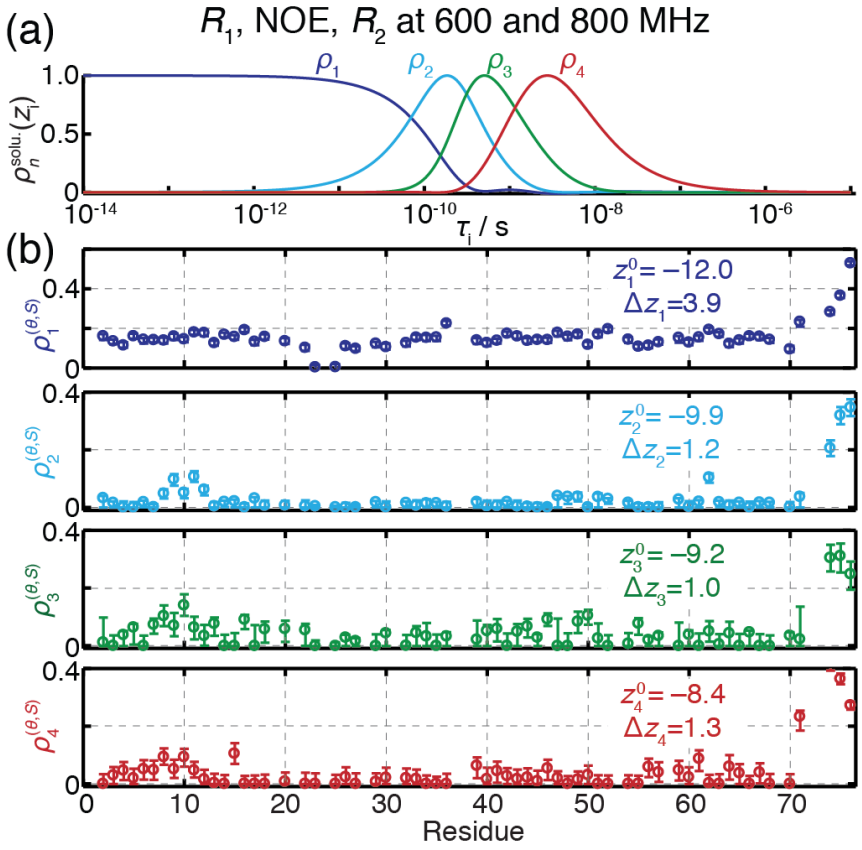

Fig. S4. Ubiquitin detector analysis using two fields (R_1, R_2, NOE at 600 and 800 MHz fields). (a) shows the detector sensitivities, (b) gives the residue-specific detector responses for each of the four detectors. Data fit is shown in Fig. S8, and detection vectors used are given in SI Table S3.

4. Model selection for dynamics detectors

In SI section 2.3, the effect of the number of detectors used on fitting and error is discussed. However, this does not tell one how many detectors is best to use. Therefore, we to try to verify that the chosen number of detectors for modeling a particular data set is the best choice, we utilize statistical model selection, via the Akaike Information Criterion (AIC), as well as several variants of this statistical test.[5] The AIC parameter is defined as

$$AIC = N\ln(\chi^2 / N) + 2K\,. \qquad \text{(S30)}$$

where χ^2 is given by

$$\chi^2 = \sum_{i=1}^{N} \frac{\left(R^i_{exper.} - R^i_{calc.}\right)^2}{\sigma_i^2}, \tag{S31}$$

and N is the total number of experiments, and here K is the number of detectors. Model selection is performed by calculating the AIC parameter and selecting the model with the smallest value.

The AIC assumes a large number of experiments so that it may be biased except in the case that $N>>K$, which is clearly not the case for NMR relaxation studies, possibly resulting in selecting a model that has too many parameters. To counter this, one may use the corrected AIC parameter (AICc),[6,7] defined as

$$AICc = N\ln(\chi^2 / N) + 2K + 2\frac{K(K+1)}{(N-K-1)}, \tag{S32}$$

but the correction term is nonetheless not always correct in the case that restrictions are placed on the fitting parameters,[8] as we do when requiring non-negative values for the detector responses. Therefore, we additionally test corrections to the AIC obtained via bootstrapping of the fit.[9] In particular, we use the AICb1 and AICb2 developed by Shang and Cavanaugh,[10] and the 632BQCV statistic developed by Bayer and Cribari-Neto.[8] We calculate the variants of the AIC parameter using data for Ubiquitin acquired at three fields (detector analysis with four detectors found in main text Fig. 9(d)). The results of the AIC tests are shown in Fig. S5.

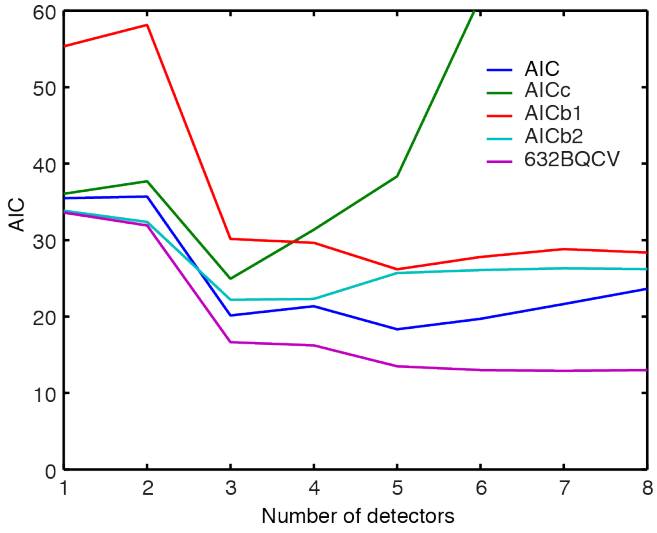

Fig. S5. Various AIC parameters as a function of the number of detectors for the analysis of backbone H–N motion in Ubiquitin. A 3 field (600, 800, 950 MHz) data set with R_1, R_2, and σ_{NH} is used. The median AIC value is reported, for all residues.

The AIC selects 5 detectors, whereas the AICc selects only 3 detectors. However, we see that the AICc rises severely for larger number of detectors, in contrast to all other tests. The bootstrap tests (AICb1, AICb2, 632BQCV) do not show a strong preference among models with 4-8 detectors. These tests should be the most reliable since they adapt the AIC correction factor to the specific model behavior based on bootstrap tests. This behavior indicates that as the model increases in complexity, the detector responses contain both more information about the internal motion, but also more noise, such that the models are ultimately of similar quality. Then for detector analysis, one could make the model selection simply based on what one considers an acceptable level of noise on the detector responses. In any case, we see that AICc analysis is not really suitable for model selection for detector analysis, and although the AIC gives similar results to the more rigorous bootstrapped tests, it is not clear that this will always be the case. Then, since AIC takes the assumption that the data set is infinitely large, it is likely better to also avoid this test.

Note that it is not straightforward to obtained a bootstrapped data set from NMR data. Typically, when performing a bootstrap, one takes the original data set, and resamples it randomly, to obtain the bootstrapped data set. However, for relaxation data, this would result in some rate constants being left out entirely, so that our detector responses are not necessarily defined for some possible bootstrapped data sets. This makes this basic approach unfeasible, so that we instead resample the *error* of our fits. Specifically, we take the initial fit to our detectors, back-calculate the rate constants, and calculate the fitting error for each experiment.

$$R^i_{exper.} = R^i_{calc.} + \epsilon_i \,, \tag{S33}$$

Here, *i* indicates an experiment of the full data set. Then, the bootstrapped data set is given for all rate constants as

$$R^i_{bootstrap.} = R^i_{calc.} + \epsilon_j \frac{\sigma_i}{\sigma_j} \,, \tag{S34}$$

where the index *j* is selected at random from all experiments in the data set with replacement, and the error is re-scaled according to the standard deviation of the experiment.

5. Model-free failure of one- and two-field data sets

As was done with a large relaxation data set in Fig. 1 (main text), it is possible to demonstrate that the model-free approach may not yield a good representation of the true motion in the case of smaller, one- and two-field data sets. We calculate relaxation rate constants here (R_1,

R_2, σ_{NH}) for motions having three correlation times, such that the correlation function is given by:

$$C(t) = \frac{1}{5}\exp(-t/\tau_r)\left[S^2 + (1-S^2)\sum_{k=1}^{3} A_k \exp(-t/\tau_k)\right], \quad \text{(S35)}$$

where $\tau_r = 4.84$ ns, and the A_i add to 1. Then, for relaxation rate constants calculated at one field (600 MHz), we fit the data to a correlation function with one internal motion (2 parameter fit):

$$C(t) = \frac{1}{5}\exp(-t/\tau_r)\left[S^2 + (1-S^2)\exp(-t/\tau_1)\right], \quad \text{(S36)}$$

or two internal motions.

$$C(t) = \frac{1}{5}\exp(-t/\tau_r)\left[S^2 + (1-S^2)(A_1\exp(-t/\tau_1) + A_2\exp(-t/\tau_2))\right], \quad \text{(S37)}$$

Here, we assume the second internal motion is sufficiently fast that it does not directly induce any relaxation, so that its value may be fixed to some arbitrarily small value ($\tau_1 = 10^{-14}$ s, three-parameter fit). Finally, when fitting data with two fields, we use the same correlation function, but allow both correlation times to vary (4 parameter fit).

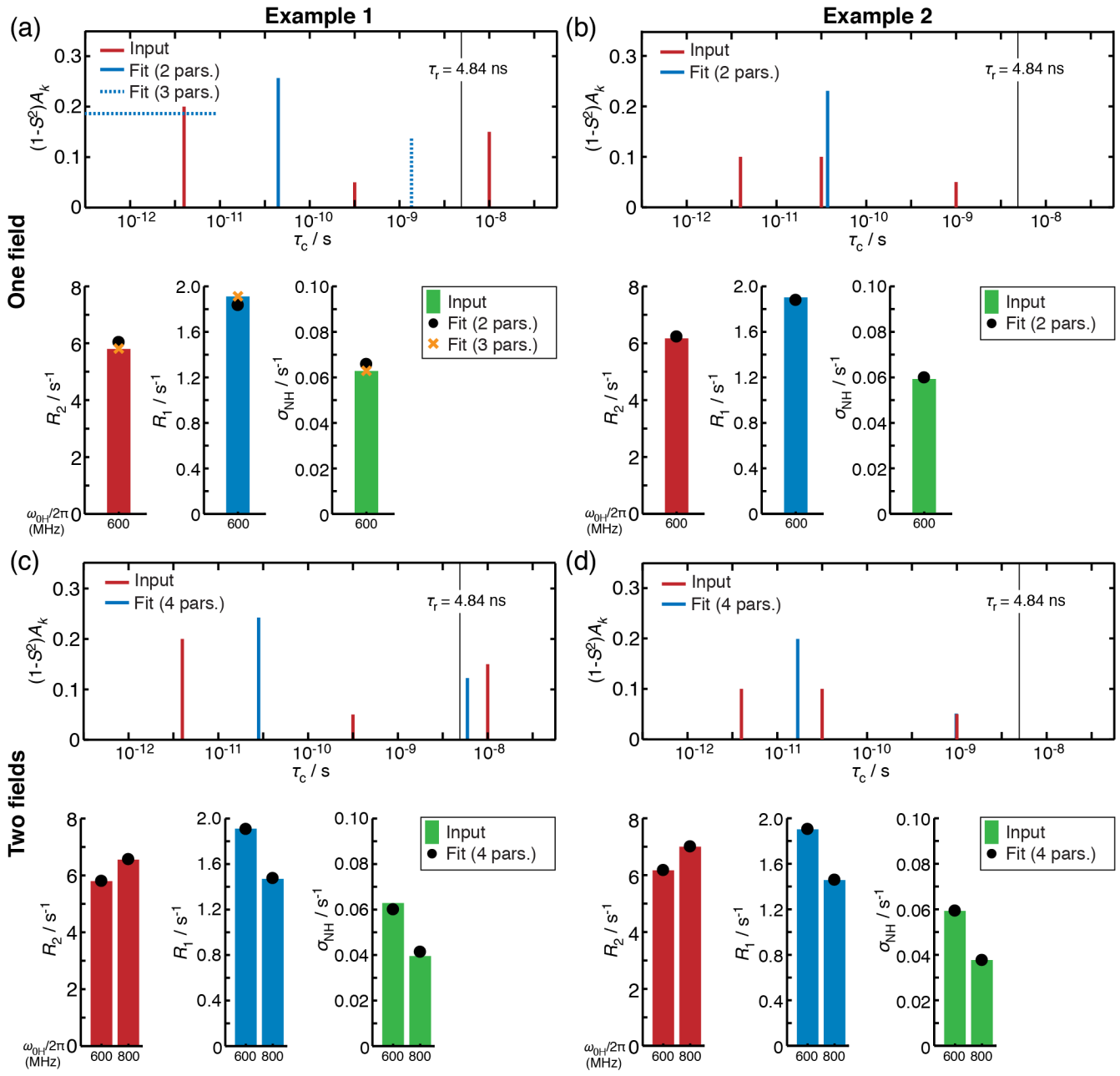

Fig. S6. Two correlation functions (example 1: (a), (c) and example 2: (b), (d)) are used to calculate R_1, R_2, and σ_{NH} rate constants at one (600 MHz: (a) and (b)) and two fields (600, 800 MHz: (c) and (d)). The correlation functions are given as line plots in the top of each subplot (red lines, giving the correlation time, τ_k, and amplitude, $(1-S^2)A_k$, of each motion), assuming a rotational correlation time of $\tau_r = 4.84$ ns The resulting rate constants are shown as colored bars in each subplot. These rate constants are then fit to models having 2, 3, or 4 free parameters (see SI Eqs. (S33) and (S34)). The resulting fit parameters are given as blue lines in the top of each subplot, and the fitted rate constants are shown as scatter points in the bottom of each subplot. In (a), the motion is fit both with a two- and three-parameter model. The results for the three-parameter model are shown as dotted lines in the top plot. Since no correlation time is fitted for the faster motion (it is fixed to $\tau_1 = 10^{-14}$ s), it is shown as a horizontal line extending from $\tau_c = 10^{-11}$ s to shorter correlation times.

We see in Fig. S6 that although the data is well-fit in all cases, the fit of the internal motion is usually far away from the input motion, as we expect when the model is simpler than

the real motion (note that in Fig. S6(d), the fitted amplitude of the shorter correlation time is approximately the sum of the amplitudes of the two shorter correlation times, and the fitted correlation time converges on the average of these two correlation times, as expected when both motions are in the extreme narrowing limit[11]). Note that in Fig. S6(a), the two parameter yields fitted rate constants that have deviated somewhat from the input, so that we also fit with three parameters, yielding an improved fit of the rate constants.

We may also investigate how well the order parameter of the internal motion is estimated. We tabulate the input and fitted order parameters. We see that the fitted order parameter is always greater than or equal to the fitted order parameter, and note that as the model complexity increases, the accuracy of the order parameter improves (assuming that using a more complex model is justified by poor fit quality of a simpler model). Such a result is expected since tumbling partially or completely masks motions with correlation times comparable to or longer than τ_r. Note that if a motion is not completely masked, then one can improve the estimation of the order parameter by using a more complex model (and including more data in the fit as necessary).

Table SI. Input vs. fitted order parameters (S^2) for each example and fit.

	Input S^2 (Ex. 1)	**Fit S^2 (Ex. 1)**	**Input S^2 (Ex. 2)**	**Fit S^2 (Ex. 2)**
1 field, 2 pars.	0.600	0.743	0.750	0.769
1 fields, 3 pars.	0.600	0.673	0.750	0.751*
2 fields, 4 pars.	0.600	0.638	0.750	0.750

*Fit not shown in Fig. S6

6. Plots of data fits

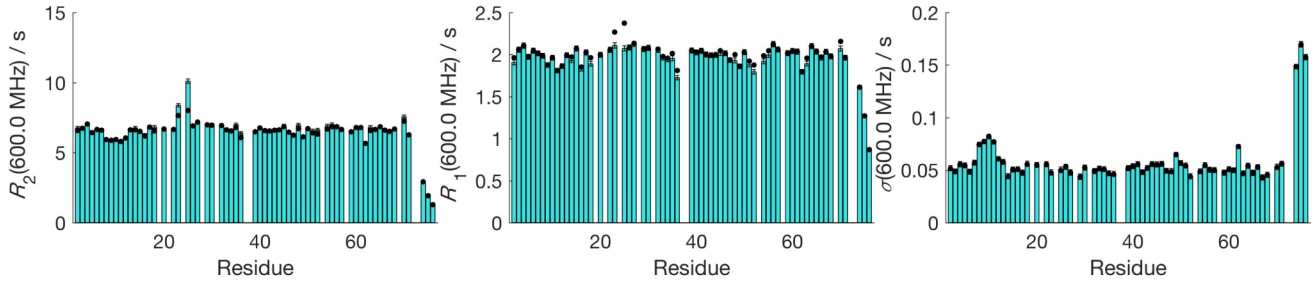

Fig. S7. Data fit of ubiquitin using only one B_0 field (from analysis shown in Fig. 7B). Each plot shows rate constants for the experiment type shown on the axis (where the field is given in parenthesis). Cyan bars give the value of the rate constant, error bars show one standard deviation, and black circles show the fitted rate constant.

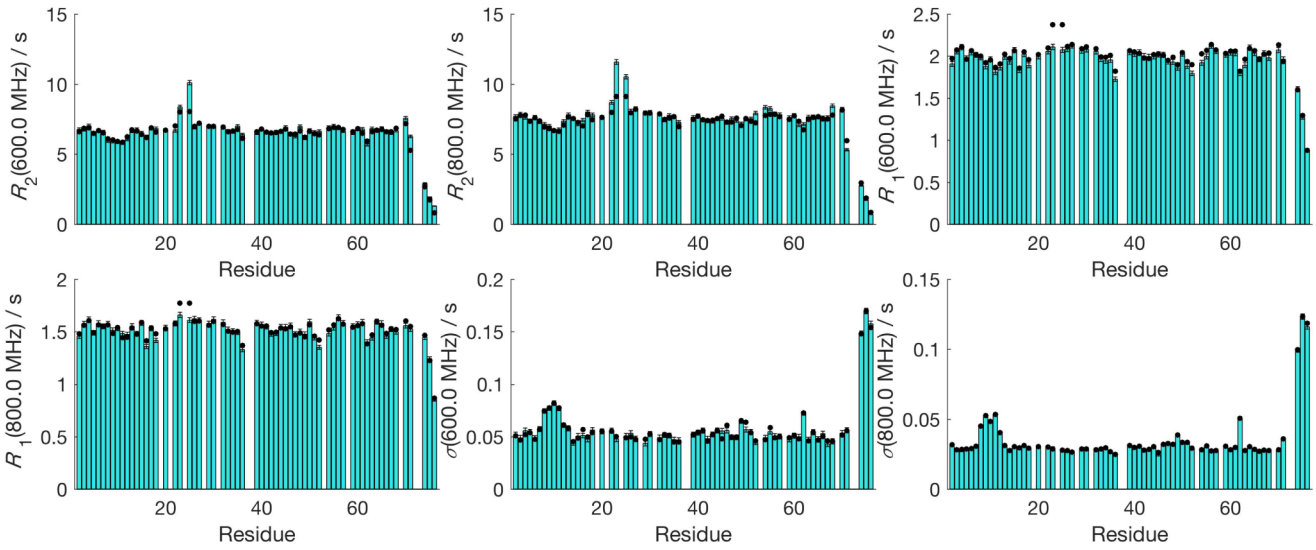

Fig. S8. Data fit of ubiquitin using two B_0 fields (from analysis shown in SI Fig. S4). Each plot shows rate constants for the experiment type shown on the axis (where the field is given in parenthesis). Cyan bars give the value of the rate constant, error bars show one standard deviation, and black circles show the fitted rate constant.

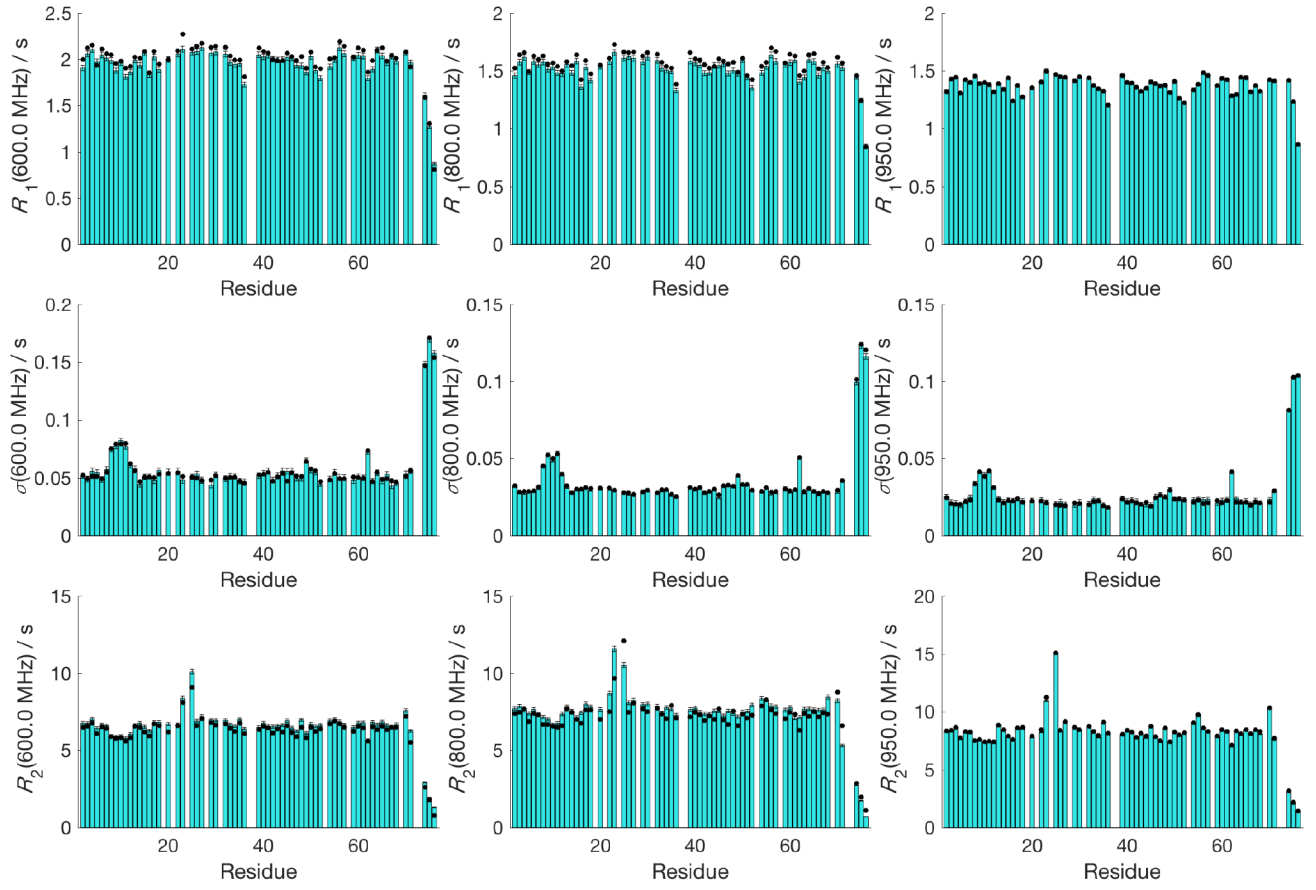

Fig. S9. Data fit of ubiquitin using three B_0 fields (from analysis shown in Fig. 7(d)). Each plot shows rate constants for the experiment type shown on the axis (where the field is given in parenthesis). Cyan bars give the value of the rate constant, error bars show one standard deviation, and black circles show the fitted rate constant.

7. Tables of detection vectors for Ubiquitin analyses

Table S2: Detection vectors for Ubiquitin analysis at one field (see Fig. 7(a)/(b))

	$\vec{r}_1$ / s^{-1}	$\vec{r}_2$ / s^{-1}	$\vec{r}_3$ / s^{-1}	R_0	B_0 / T
$R_{2,600}$	-8.038	-6.592	-2.736	8.047	14.1
$R_{1,600}$	-2.381	-1.347	0.662	2.382	14.1
$\sigma_{HN,600}$	-0.050	0.346	0.021	0.050	14.1

Other parameters: δ_{HN}=-22945 Hz, $\Delta\sigma_N$=169.5 ppm, τ_r=4.84 ns

Table S3: Detection vectors for Ubiquitin analysis at two fields (see Fig. S4)

	$\vec{r}_1$ / s^{-1}	$\vec{r}_2$ / s^{-1}	$\vec{r}_3$ / s^{-1}	$\vec{r}_4$ / s^{-1}	R_0	B_0 / T
$R_{2,600}$	-8.037	-3.675	-3.701	-2.983	8.047	14.1
$R_{1,600}$	-9.123	-4.271	-4.213	-3.542	2.382	14.1
$\sigma_{HN,600}$	-2.386	-0.640	-0.793	0.632	0.050	14.1
$R_{2,800}$	-1.776	-0.390	-0.322	0.874	9.140	18.8
$R_{1,800}$	-0.051	0.228	0.186	0.023	1.781	18.8
$\sigma_{HN,800}$	-0.028	0.234	0.078	0.019	0.028	18.8

Other parameters: δ_{HN}=-22945 Hz, $\Delta\sigma_N$=169.5 ppm, τ_r=4.84 ns

Table S4: Detection vectors for Ubiquitin analysis at three fields (see Fig. 7(c)/(d))

	$\vec{r}_1$ / s^{-1}	$\vec{r}_2$ / s^{-1}	$\vec{r}_3$ / s^{-1}	$\vec{r}_4$ / s^{-1}	$\vec{r}_{ex}$ / s^{-1}	R_0	B_0 / T
$R_{2,600.3}$	-8.040	-2.551	-3.772	-3.442	1	8.048	14.1
$R_{2,800.4}$	-9.129	-2.972	-4.318	-4.059	1.778	9.142	18.8
$R_{2,949.4}$	-10.298	-3.427	-4.857	-4.722	2.507	10.317	22.3
$R_{1,600.3}$	-2.390	-0.317	-0.962	0.618	0	2.381	14.1
$R_{1,800.4}$	-1.783	-0.133	-0.540	0.944	0	1.790	18.8
$R_{1,949.4}$	-1.521	-0.074	-0.264	1.062	0	1.523	22.3
$\sigma_{HN,600.3}$	-0.050	0.182	0.254	0.038	0	0.050	14.1
$\sigma_{HN,800.4}$	-0.029	0.196	0.148	0.021	0	0.028	18.8
$\sigma_{HN,949.4}$	-0.020	0.194	0.097	0.017	0	0.020	22.3

Other parameters: δ_{HN}=-22945 Hz, $\Delta\sigma_N$=169.5 ppm, τ_r=4.84 ns

References

1 A. A. Smith, M. Ernst, and B. H. Meier, J. Chem. Phys. **148** (4), 045104 (2018).

2 G. Golub and W. Kahan, Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis **2** (2), 205 (1965).

3 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).

[4] A. A. Smith, M. Ernst, and B. H. Meier, Distortion Free Relaxation Analysis Technique software. URL: http://difrate.sourceforge.net, (2018).
[5] H. Akaike, IEEE T. Automat. Contr. **19** (6), 716 (1974).
[6] C. M. Hurvich and C.-L. Tsai, J. Time Series. Anal. **14** (3), 271 (1993).
[7] N. Sugiura, Commun. Stat.-Theor. M. **7**, 13 (1978).
[8] F. Bayer and F. Cribari-Neto, arXiv:1405.4525 [stat.CO] **24** (4), 776 (2010).
[9] B. Efron, The Annals of Statistics **7** (1), 1 (1979).
[10] J. Shang and J. Cavanaugh, Comput. Stat. Data. An. **52** (4), 2004 (2008).
[11] G. Lipari and A. Szabo, J. Am. Chem. Soc. **104** (17), 4546 (1982).