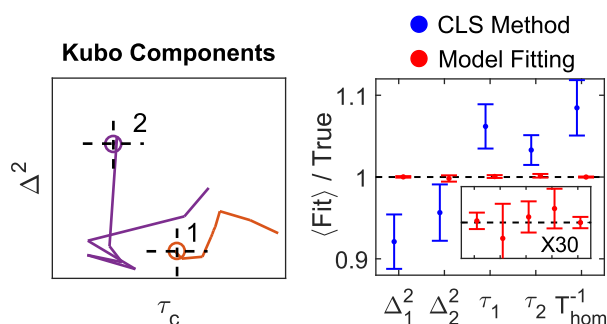# Least-Squares Fitting of Multidimensional Spectra to Kubo Lineshape Models

Kevin C. Robben, Christopher M. Cheatum*

Department of Chemistry, University of Iowa, Iowa City, Iowa 52242, USA

TOC Graphic

Abstract

We report a comprehensive study of the efficacy of least-squares fitting of multidimensional spectra to generalized Kubo lineshape models and introduce a novel least-squares fitting metric, termed the Scale Invariant Gradient Norm (SIGN), that enables a highly reliable and versatile algorithm. The precision of dephasing parameters is between 8× to 50× better for nonlinear model fitting compared to the CLS method, which effectively increases data acquisition efficiency by one to two orders of magnitude. Whereas the center-line-slope (CLS) method requires sequential fitting of both the nonlinear and linear spectra, our model fitting algorithm only requires nonlinear spectra, but accurately predicts the linear spectrum. We show an experimental example in which the CLS time constants differ by 60% for independent measurements of the same system, while the Kubo time constants differ by only 10% for model fitting. This suggests that model fitting is a far more robust method of measuring spectral diffusion than the CLS method, which is more susceptible to structured residual signals that are not removable by pure solvent subtraction. Statistical analysis of the CLS method reveals a fundamental oversight in accounting for the propagation of uncertainty by Kubo time constants in the process of fitting to the linear absorption spectrum. A standalone desktop app and source code for the least-squares fitting algorithm are freely available with example lineshape models and data. We have written the MATLAB source code in a generic framework where users may supply custom lineshape models. Using this application, a standard desktop fits a 12-parameter generalized Kubo model to a 10^6 data-point spectrum in a few minutes.

Introduction

The Kubo lineshape is a common model of spectral diffusion in frequency fluctuation correlation functions (FFCFs) owing to its simple, closed-form expression and flexibility to describe the limiting cases of homogeneous and inhomogeneous dephasing.[1] As the FFCF is not a direct experimental observable, a variety of approximate metrics have been used to extract an approximate FFCF from 2D frequency resolved line shapes in two-dimensional infrared (2D IR) spectra. Some of these approaches include the nodal-line slope,[2-4] dynamic linewidth,[5] ellipticity,[6, 7] covariance,[8] spectral phase slope,[9] inhomogeneity index,[10] eccentricity,[11] center line slope (CLS),[12] inverse center line slope (invCLS),[13, 14] and the correlation coefficient obtained by a 2D Gaussian fit.[15] The most popular among these has been the CLS, which, relative to some other metrics, is more reliable at lower SNR, is invariant to lineshape interference from anharmonic peaks or phase twist, and is invariant to apodization time for time constants and *relative* amplitudes.[16]

There are, however, several shortcomings of the CLS method, which are also characteristics of most of the other common metrics. First, it is unreliable for characterizing relatively fast processes due to the short-time approximation,[12, 13] which, following second-order cumulant expansion, ignores dephasing during coherence times. Second, it requires a second step of fitting to the linear absorbance spectrum to obtained absolute values for Kubo amplitudes and homogeneous dephasing. This fitting is problematic in situations where the linear absorption is either unavailable or unreliable due to dilute reporters, weak extinction coefficients, or spectral congestion. Furthermore, linear absorption spectra are often inaccurate near the baseline, which can distort the resulting fit parameters. Third, *absolute* values of Kubo amplitudes and homogeneous dephasing are sensitive to the apodization time, meaning results may vary depending on the duration of coherence time measured or how quickly the apodization filter tapers to zero.[16] Fourth, the CLS method still requires a remarkably high SNR (e.g. ~100:1) to yield reliable results.

Model fitting to the nonlinear waiting-time-dependent spectra provides a natural solution to these issues. Directly fitting the data to a user-supplied lineshape model does not depend on the short-time approximation. It allows users to match the time/frequency domain(s) in which they collect data, and thereby mitigate apodization bias. It does not depend on fitting the linear absorption spectrum, and therefore, enables accurate measurements of spectral diffusion in a variety of scenarios that were previously inaccessible. Finally, as we will show, model fitting is reliable at far lower SNR (e.g. 10:1), is far less susceptible to structured residual signals than the CLS method, and yields reliable uncertainties for the measured lineshape parameters.

This manuscript provides a comprehensive study of how fitting generalized Kubo lineshapes to multidimensional spectra compares to the CLS method in terms of the accuracy, precision and reliability of the resulting parameters. While Garret-Roe and coworkers have shown several examples of model fitting 2D IR waiting-time series using the `fmincon` function in MATLAB,[17-20] a comparison of accuracy between model fitting and the CLS method by fitting to simulated spectra, where true values of parameters are known, has been missing. We find our model fitting routine improves precision over the CLS method by 8× to 15× on average for

Kubo time constants and 8× to 50× for Kubo amplitudes and homogeneous dephasing, which is due in part to a novel figure of merit used in our fitting algorithm that we refer to as the scale invariant gradient norm (SIGN). We find that numerical instabilities associated with some fitting parameters appearing nearly indistinguishable at certain points in parameter space are the primary cause of sudden ceasing, which could be mistaken for local minima. Importantly, the SIGN readily identifies these events, which enables swift correction by our fitting algorithm.

We begin the manuscript by describing our approach, including a description of the least-squares fitting program, estimation of error, the introduction of the scale invariant gradient norm, a brief review of multicollinearity (or ill-conditioning) in fitting problems, preprocessing of data prior to fitting, and description of hardware and software used in measurements and least-squares fitting. We then examine several experiments including a side-by-side comparison of model fitting and the CLS method for 100 trials of simulated data, fitting with too many or too few Kubo components, fitting to low SNR data, fitting to data with phasing errors, fitting to experimental data, and fitting to undersampled data. We then conclude with a discussion of recommended practices for model fitting.

Methods and Materials

A. Least-Squares Fitting Algorithm

The Gauss-Newton algorithm is a common approach for model fitting.[21] Within the least-squares routine, experimentally measured data (provided as a multidimensional input) are concatenated into a one-dimensional vector $\mathbf{D}$ ($N_D \times 1$) where information regarding dimensionality is preserved in the ordering of data. Throughout the text we use bold letters to denote vectors and matrices. We denote the residual between data $\mathbf{D}$ and lineshape model $\mathbf{M}(\mathbf{p})$ by $\mathbf{r}$ in Eq. 1 where $\mathbf{p}$ ($N_p \times 1$) is a vector of variable fitting parameters. We define "parameter" as any number subject to change with the measured system (e.g. the center frequency, homogeneous lifetime, etc.), while the preceding adjectives "constant" or "variable" refer to the status of a parameter during least-squares fitting. For generalized least-squares, the cost function $C(\mathbf{p})$ (a.k.a. χ^2) is equal to the quadratic form in Eq. 2 where $\mathbf{V_D}$ ($N_D \times N_D$) is proportional to the data variance-covariance matrix,[22] and superscript $\mathbf{T}$ denotes the transpose. In the simplest case of uniform and uncorrelated noise, $\mathbf{V_D}$ is equal to the identity matrix and the cost function $C(\mathbf{p}) = |\mathbf{r}|^2$. For more complicated cases of noise, a detailed discussion of $\mathbf{V_D}$ is provided section B below.

$$\mathbf{r} = \mathbf{D} - \mathbf{M}(\mathbf{p}) \tag{Eq. 1}$$

$$C(\mathbf{p}) = \mathbf{r^T V_D^{-1} r} \tag{Eq. 2}$$

The objective of least-squares fitting is to minimize Eq. 2 subject to $\mathbf{p}$. Because $\mathbf{r}$ depends nonlinearly on $\mathbf{p}$, minimizing the cost function $C(\mathbf{p})$ requires an iterative process: $\mathbf{p_{i+1}} = \mathbf{p_i} + \mathbf{\Delta p}$. At each iteration, the second order Taylor series shown in Eq. 3 locally approximates $C(\mathbf{p})$ where ∇C ($1 \times N_p$) and $\nabla\nabla C$ ($N_p \times N_p$) are the gradient and Hessian of $C(\mathbf{p})$ with respect to $\mathbf{p}$.

$$C(\mathbf{p} + \Delta\mathbf{p}) \approx C(\mathbf{p}) + \nabla C \Delta\mathbf{p} + \frac{1}{2} \Delta\mathbf{p}^T \nabla\nabla C \Delta\mathbf{p} \qquad\qquad \text{Eq. 3}$$

Applying the gradient ∇ to Eq. 2, we find a more useful expression for ∇C in Eq. 4 given in terms of $\mathbf{r}$, $\mathbf{V_D}$, and the Jacobian $\mathbf{J}$ ($N_D \times N_p$) which is a matrix composed of partial derivatives $\partial\mathbf{M}/\partial p_k$ computed by finite-difference approximation.

$$\nabla C = -2\mathbf{r}^T \mathbf{V_D}^{-1} \mathbf{J} \qquad\qquad \text{Eq. 4}$$

Eq. 5 provides an expression for $\nabla\nabla C$ in terms of the residual-weighted Hessian of the model ($\mathbf{H}_{j,k} = \mathbf{r}^T \mathbf{V_D}^{-1} \partial^2\mathbf{M}/\partial p_j \partial p_k$). Simply put, the most salient difference between several popular algorithms is in how they compute $\nabla\nabla C$. Newton's algorithm computes it exactly as $\nabla\nabla C = 2(\mathbf{J}^T\mathbf{V_D}^{-1}\mathbf{J} - \mathbf{H})$.[21] Gauss-Newton computes it approximately as $\nabla\nabla C = 2\mathbf{J}^T\mathbf{V_D}^{-1}\mathbf{J}$. Levenberg-Marquardt computes a more stable approximation $\nabla\nabla C = 2(\mathbf{J}^T\mathbf{V_D}^{-1}\mathbf{J} + \lambda\mathbb{1})$,[21, 23, 24] or some variation thereof,[25] where λ is known as the damping parameter and $\mathbb{1}$ is the identity matrix. Finally, Steepest Descent simply assumes $\nabla\nabla C = \mathbb{1}$.[21] The advantage of the Gauss-Newton algorithm over Newton is time saved in not computing $\mathbf{H}$, which is usually quite significant. Newton is also susceptible to convergence problems far from the global minimum, unlike the other algorithms mentioned. The Gauss-Newton approximation is often justified because near the global minimum the residual $\mathbf{r}$ is relatively small and mostly random with zero mean, implying $\mathbf{H}$ is negligible. The Levenberg–Marquardt algorithm interpolates between the limiting cases of Gauss-Newton ($\lambda \to 0$) and Steepest Descent ($\lambda \to \infty$). The advantage of Levenberg–Marquardt is the added stability of inverting $\nabla\nabla C$ due to the $\lambda\mathbb{1}$ term, however, this is irrelevant in our case as our algorithm guards against singular $\nabla\nabla C$ (Section C below). The disadvantage of Levenberg–Marquardt is that successful optimization for λ is difficult to predict and usually requires a dynamic routine. Furthermore, as λ increases, Levenberg–Marquardt behaves more like Steepest Descent which is slower to converge near minima because $\nabla\nabla C \to \lambda\mathbb{1}$ ignores the true curvature of $C(\mathbf{p})$. Therefore, we have chosen the Gauss-Newton approximation in Eq. 5.

$$\nabla\nabla C = 2(\mathbf{J}^T\mathbf{V_D}^{-1}\mathbf{J} - \mathbf{H}) \approx 2\mathbf{J}^T\mathbf{V_D}^{-1}\mathbf{J} \qquad\qquad \text{Eq. 5}$$

The minimum of Eq. 3, $\Delta\mathbf{p}$, is obtained using the MATLAB syntax $\Delta\mathbf{p} = -\nabla\nabla C \backslash \nabla C^T$ where "\" corresponds to the MATLAB function `mldivide`, which solves the linear system of equations in Eq. 6. Note that $\nabla\nabla C$ may not be invertible on occasion. As is standard practice with nonlinear fitting routines, $\Delta\mathbf{p}$ undergoes a quality check at the end of each iteration to ensure the move is productive. In particular, the program uses a backtracking line search subject to the Armijo condition. And finally, the program compares the new position $\mathbf{p}_{i+1} = \mathbf{p}_i + \Delta\mathbf{p}$ to the parameter boundaries provided by the user and corrects $\Delta\mathbf{p}$ if necessary.

$$\nabla\nabla C \cdot \Delta\mathbf{p} = -\nabla C^T \qquad\qquad \text{Eq. 6}$$

Occasions may arise in which the solution for $\Delta\mathbf{p}$ in Eq. 6 is inaccurate or unacceptable in a directional sense. Consequently, iterative changes in $\Delta\mathbf{p}$ approach zero even though ∇C is clearly nonzero. We refer to this as algorithmic stalling. Stalling is a separate issue from a local minimum in that $\nabla C = 0$ in a local minimum. We find that simply sending $\mathbf{p}$ to a random point

within user supplied boundaries, which we refer to as a *random restart,* is a reliable strategy for resolving a stall. This approach is closely related to *multistart,*[26] which is a shotgun strategy for problems plagued by local minima.

B. Uncertainty of Fit and Nonuniform Noise

The parameter variance-covariance matrix $\mathbf{V_p}$ ($N_p \times N_p$) given by Eq. 7 provides the uncertainty of the variable fitting parameters assuming that $\mathbf{p}$ is located at the global minimum.[22] We provide a derivation of Eq. 7 in SI section D.

$$\mathbf{V_p} = \frac{C(\mathbf{p})}{N_D - N_p} \left(\mathbf{J}^T\mathbf{V_D^{-1}J}\right)^{-1} \qquad\qquad \text{Eq. 7}$$

Note that Eq. 7 assumes $\mathbf{V_D}$ ($N_D \times N_D$) is proportional to the data variance-covariance matrix. If noise is uncorrelated across $\mathbf{D}$, but not necessarily uniform, then $\mathbf{V_D}$ is a diagonal matrix with elements proportional to the variance of each datum. For example, if datum i is averaged twice as much as datum k, then $\mathbf{V_D^{-1}}_{ii}/\mathbf{V_D^{-1}}_{kk} = 2/1$, and hence, the convenience of referring to $\mathbf{V_D^{-1}}$ instead of $\mathbf{V_D}$. We emphasize that $\mathbf{V_D}$ need only be *proportional* to the true variance-covariance matrix in Eq. 7 since $C(\mathbf{p}) \propto \mathbf{V_D^{-1}}$ and $\left(\mathbf{J}^T\mathbf{V_D^{-1}J}\right)^{-1} \propto \mathbf{V_D}$. If noise is both uncorrelated and uniform across $\mathbf{D}$, then $\mathbf{V_D^{-1}}$ is the identity matrix.

Accounting for correlated noise more generally is challenging because the nondiagonal data variance-covariance matrix $\mathbf{V_D}$ of size $N_D \times N_D$ might easily occupy a terabyte of memory for multidimensional spectra. Hence, for feasibility sake, our program assumes $\mathbf{V_D}$ is diagonal which is typical for most other fitting programs. Consequently, the fitting algorithm is most optimal for spectra with uncorrelated noise. While conventional referencing schemes are unreliable for achieving uncorrelated noise, *calibrated* referencing schemes are known to achieve virtually uncorrelated noise,[27-29] which can also be realized using 100 kHz Yb laser systems.[30-32] Nevertheless, SI section F provides a comparison of model fitting to edge-pixel referenced[27] and unreferenced data, which suggests that model fitting to unreferenced data is still reliable setting aside the expected gain in uncertainty from the larger noise.

C. Scale Invariant Gradient Norm $|\widetilde{\nabla C}|$

Stopping criteria are a notorious complication with fitting algorithms. They are often based on user specified thresholds. Three common examples are $|\Delta\mathbf{p}| < 10^{-4}$, $|C(\mathbf{p_{i+1}}) - C(\mathbf{p_i})| < 10^{-5}$ or $|\nabla C| < 10^{-6}$. The threshold values of 10^{-4}, 10^{-5} and 10^{-6} in these examples are arbitrary and may strongly depend on factors such as the scaling of parameters $|\mathbf{p}|$, data $|\mathbf{D}|$, noise $|\mathbf{V_D}|$, and number of data points N_D. Therefore, users must reconsider these thresholds on a case-by-case basis, often empirically. To that end, we propose a new stopping criterion which we refer to as the scale invariant gradient norm, $|\widetilde{\nabla C}|$, defined in Eq. 8. We motivate this expression by unit analysis of $|\nabla C|$: $C(\mathbf{p})$ cancels with the numerator of ∇C and each element of $\sqrt{\text{diag}(\mathbf{V_p})}$ cancels with a corresponding element of $\partial\mathbf{p_i}$ in the denominator of ∇C. In Eq. 8, the numerator is

evaluated at iteration *i*, where $\mathbf{\nabla C}$ is of size $1 \times N_p$ (Eq. 4) and $\sqrt{\text{diag}(\mathbf{V_p})}$ is of size $N_p \times 1$ (the diagonal root of Eq. 7), while $C(\mathbf{p})$ (a scalar, Eq. 2) in the denominator is evaluated at iteration *i*-1 (evaluating at iteration *i*-1 is better behaved during random restarts).

$$|\widetilde{\nabla C}| = \frac{\left| \mathbf{\nabla C} \sqrt{\text{diag}(\mathbf{V_p})} \right|_i}{C_{i-1}(\mathbf{p})} \qquad \text{Eq. 8}$$

In addition to serving as a stopping criterion, $|\widetilde{\nabla C}|$ reliably indicates when the algorithm is stalled, which is resolved by random restart. Our stopping criteria require that the previous three iterations have $|\widetilde{\nabla C}| < 10^{-9}$ and have $C(\mathbf{p})$ within 10% of the lowest $C(\mathbf{p})$ encountered in all previous iterations, which provides some moderate protection against local minima. Regardless, we show that local minima are virtually nonexistent for three-level systems, so the second stopping criterion is somewhat moot. If an iteration does not meet stopping criteria, then the program checks for a stall. Our stalling criteria require that the last three iterations have less than 1% deviation in $|\widetilde{\nabla C}|$ and less than 1% deviation in $C(\mathbf{p})$. The program also triggers a stall if $\mathbf{\nabla\nabla C}$ is singular, or nearly singular. If no stall is detected, then the program continues to the next iteration.

D. Multicollinearity Considerations

Considerations of multicollinearity, also known as ill-conditioning, are *essential* to developing a successful fitting model. We speculate that the limited application of multidimensional fitting algorithms in nonlinear spectroscopy to date is due in part to overlooking this aspect of models. A fitting problem is said to be multicollinear (or ill-conditioned) if column vectors of the Jacobian $\mathbf{J}$ are nearly linearly dependent, which causes instability in computing inverse matrices associated with $\mathbf{J}$, such as the Hessian $\mathbf{\nabla\nabla C}^{-1}$ and $\mathbf{V_p}$.[33, 34] Consequently, the calculation of $\mathbf{\Delta p}$ in Eq. 6 may be inaccurate or unstable (i.e. a major cause of stalling) and covariances among those nearly linearly dependent parameters become overwhelming. In more extreme cases of multicollinearity, tiny perturbations of noise cause wild fluctuations in fitting parameters,[35, 36] and error estimates become useless.[33, 35, 37]

Not surprisingly, parameter redundancy, or indistinguishability between parameters, drives multicollinearity[38] and should be avoided whenever possible. Here we list a few examples of multicollinearity.

Example 1: Homogeneous dephasing is a limiting case of a Kubo lineshape where T_{hom}^{-1} approaches the product $\Delta^2\tau$. In this case, Δ^2 and τ are indistinguishable. Hence, homogeneous dephasing is modeled by a single fitting variable, T_{hom}^{-1}.

Example 2: We refer to a pair of Kubo components with similar correlation times ($\tau_1 \approx \tau_2$), but different amplitudes ($\Delta_1^2 \neq \Delta_2^2$) as *degenerate*. In this case, Δ_1^2 and Δ_2^2 become indistinguishable due to their linear dependence in the FFCF, corresponding to $\Delta_1^2 \exp(-t/\tau_1) + \Delta_2^2 \exp(-t/\tau_2) \approx (\Delta_1^2 + \Delta_2^2) \exp(-t/\tau_1)$.

Example 3: When modeling an isotropic response, the total homogeneous dephasing (T_{Hom}) has three contributions from pure dephasing (T_2^*), vibrational lifetime (T_{LT}), and orientational relaxation (T_{or}) : $1/T_{hom} = 1/T_2^* + 1/2T_{LT} + 1/3T_{or}$. If the waiting-time axis is sufficiently well-sampled to ensure linear independence of T_{LT}, the remaining three lifetimes T_{hom}, T_2^* and T_{or} are still multicollinear with one another. This is equivalent to trying to solve an algebraic equation for the values of x, y and z given only that $x = y + 5/2 + z/3$. There are infinite possible answers. So, we reduce the number of unknowns to simply T_{Hom} and T_{LT}, without specifying T_2^* and T_{or} while fitting the isotropic response but enforce the boundary condition $1/T_{hom} > 1/2T_{LT}$.

Example 4: Due to calibration error, the 0-1 peak may slightly differ in location on the pump and probe axes (e.g. <1 cm⁻¹). Therefore, it is okay to model a calibration error along one of the axes, but modeling it along both axes, or modeling two calibration errors (one for each axis), would cause indistinguishability between the calibration error(s), the 0-1 center frequency and the anharmonic shift.

A common measure of multicollinearity is the Variance Inflation Factor (VIF).[34, 39] As the name suggests, the VIF is the factor by which the variance of a parameter inflates due to collinearity with other variable parameters. The VIF of the ith parameter may be computed empirically by measuring the variance over many simulated trials for two scenarios: (1) one in which all fitting parameters are varied during each fit (as is usual) and (2) only the ith parameter is varied during the fit and all other fitting parameters are held constant at their true values. Then the VIF is the ratio of the former to the latter. However, the empirical method may be time consuming, so we elect for the equivalent theoretical expression: Column vectors of the Jacobian $\mathbf{J}$ are normalized to $\hat{\mathbf{J}}$ such that all diagonal elements of $\hat{\mathbf{J}}^T \mathbf{V_D^{-1}} \hat{\mathbf{J}}$ are equal to one. Then the VIF of the ith parameter is equal to the ith diagonal element of $(\hat{\mathbf{J}}^T \mathbf{V_D^{-1}} \hat{\mathbf{J}})^{-1}$.[34, 39] To understand this, consider the limiting case of perfectly orthogonal model parameters and uniform, uncorrelated noise (i.e. $\mathbf{V_D}$ is the identity matrix), then $\hat{\mathbf{J}}^T \mathbf{V_D^{-1}} \hat{\mathbf{J}}$ is the identity matrix, and hence, VIF is equal to 1 for every parameter. In the other limiting case in which any two or more parameters are perfectly linearly dependent, then $\hat{\mathbf{J}}^T \mathbf{V_D^{-1}} \hat{\mathbf{J}}$ is rank deficient, causing $\det(\hat{\mathbf{J}}^T \mathbf{V_D^{-1}} \hat{\mathbf{J}}) = 0$, and the diagonal of $(\hat{\mathbf{J}}^T \mathbf{V_D^{-1}} \hat{\mathbf{J}})^{-1}$ blows up to infinity. For that reason, we recommend the true matrix inverse for testing VIF, not a pseudo inverse.

E. Preprocessing Prior to Fitting

We prefer to fit data in the original measurement domain, which is (τ_1, T_w, ω_3). Therefore, our model requires a Fast Fourier Transform (FFT) along the probe axis. The FFT assumes equal spacing along the probe axis, but, due to the spectrograph, our experimental data points are nonlinearly spaced along the probe axis. Therefore, to match the data and model in Eq. 1, users should preprocess the data by interpolating along a linear probe axis. This is a straightforward task using the `spline` function in MATLAB. The GUI version of our program does this automatically when loading data.

Time domain data are commonly padded by an equal number of zeros just prior to the Fourier Transform to enforce causality, which we call *causal* zero-padding. We refer to zero padding beyond this point as *superfluous*, but it is a common practice to interpolate the data in the frequency domain. More precisely, this approach results in sinc interpolation in the frequency domain. Sinc ringing is nearly always mitigated by choosing a non-rectangular windowing function to make the data smoothly taper to zero prior to zero padding.[40]

Both superfluous zero padding and apodization manipulate data in different ways, and these perturbations will propagate into the fitting results. In fact, assuming the true parameters are known, the change in fitting parameters **dp** due to a perturbation in data **dD** is computed by Eq. S10. Therefore, we generally recommend fitting data in the original measurement domain without causal or superfluous zero padding, or apodization. When this is not an option, the data should be transformed back into the original measurement domain and the non-rectangular apodization window inverted to retrieve the original data. Careful consideration is needed when constructing the inverse window function (1) to avoid divide by zero or near divide by zero and (2) whether the original filter had ½ scaling of the DC component, which is common practice.[40, 41]

Data are collected using a pulse shaper[42, 43] which ensures accurate phasing of all 2D IR spectra, though we have added an optional fitting parameter to account for a uniform zero order phasing error across all spectra. Results of model fitting to simulated data with phasing errors are provided in supporting information section G.

F. Computer, Software and Computational Time

We use a standard laptop to run model fitting and data analysis in MATLAB R2020a i.e. an Intel Core i7-8550U CPU @ 1.80GHz, 16 GB of RAM and (optionally) an NVIDIA GeForce GTX 1050 GPU, 4 GB GDDR5. The standalone desktop app (available for Windows and Mac users), MATLAB source code, and experimental data are freely available at https://github.com/kevin-robben/model-fitting. Detailed instructions for reproducing all data and analyses are provided in section H of the supporting information.

The time needed to fit a single waiting-time series is approximately equal to $\zeta \times N_D \times N_p$ where ζ is a constant specific to the computer, N_D is the number of data points and N_p is the number of variable fitting parameters. The Jacobian **J** requires $2N_p$ unique calculations of the model **M(p)** for the central finite-difference approximation, and hence, the bottle neck of the program is computing **M(p)**. We estimate $\zeta \approx 1.3 \times 10^{-7}$ minute/point/parameter for CPU computing on the laptop described above. For spectra smaller than 10^6 points, ζ is roughly the same between CPU and GPU computing. For spectra larger than 2×10^6 points, GPU computing reduces ζ by a factor of two or greater.

Aside from GPU computing, computational time is reduced by fitting to fewer data points, which may be achieved by averaging more laser shots with the minimum necessary data points. For a simple three level system, the pump axis can be shortened to e.g. 16 data points with careful consideration of the rotating frame. This may be achieved by measuring just 16 points

along the pump axis with e.g. 250 fs steps, or alternatively, by deleting excess data points along the pump axis in the frequency domain that fall outside the region of interest, and then transforming back to the time domain. However, as mentioned above in section E, careful consideration of the window function is needed when switching between time and frequency domains.

G. Linear and 2D IR Measurements

We collect FTIR measurement on a Bruker Tensor 27 with 1 cm^{-1} resolution. 2D IR measurements are collected at 2 kHz with ~150 fs pulses, centered at ~2150 cm^{-1} and magic-angle polarization.[27] The 2020 data are collected with 15 μJ pump energy and 200 mM MeSCN in DMSO. The 2021 data are collected with 2 μJ pump energy and 400 mM MeSCN in DMSO. Edge-pixel referencing subtracts correlated local-oscillator noise.[27] The 2020 data are collected with a 4-pulse, real-valued phase cycle while the 2021 data are collected with an 8-pulse, complex-valued phase cycle. Comparative tests have led us to conclude that model fitting works equally well for fitting to real-valued or complex-valued free induction decays, given an equal number of laser shots.

H. Lineshape Models

Data are modeled as the isotropic response of a three-level system with $\text{FFCF}(t) = \frac{\delta(t)}{T_{\text{Hom}}} + \sum_i \Delta^2 \exp(-t/\tau_i)$ where the first term accounts for homogeneous dephasing and the summation accounts for multiple Kubo components when applicable. Homogeneous dephasing of the 1-2 transition is modified during the second coherence time to account for lifetime broadening of the 1-2 transition.[44] Further details and complete equations are provided in supporting information section C.

Results and Discussion

A. Fitting Simulated Data

First, we test the accuracy and precision of the algorithm by model fitting to simulated data and comparing the fit parameters to known, true values. We simulate the isotropic response of the C≡N stretch of MeSCN in H_2O as characterized by Yuan and Fayer.[45-47] We add Gaussian noise (SNR ~600:1) to the free induction decay to simulate experimental data. Here we define signal in the SNR calculation as the peak magnitude of the 0-1 transition of the transient absorption spectrum at zero waiting time. Our choice to simulate data with 600:1 SNR avoids occasional unphysical results that would otherwise occur in the CLS analysis of data with lower SNR over the course of 100 trials. Nevertheless, we also show that model fitting can be reliable for data with SNR of 10:1 (vide infra) subject to the expected increase in variance predicted by Eq. 7.

The variable fitting parameters comprising **p** are as follows: (1) the 0-1 peak amplitude A_{01} (2) the 1-2 peak amplitude A_{12} (3) the 0-1 center frequency ω_{01} (4) a calibration mismatch error between the pump and probe axes $\delta\omega_1$ (5) the anharmonicity Δ_{Anh} (6) the 1st Kubo time constant τ_1 and (7) squared amplitude Δ_1^2 (8) the 2nd Kubo time constant τ_2 and (9) squared amplitude Δ_2^2

(10) a scaling factor for the Kubo amplitude of the 1-2 transition relative to the 0-1 β (i.e. $\Delta_1^2(1\text{-}2) = \beta^2 \Delta_1^2(0\text{-}1)$ and $\Delta_2^2(1\text{-}2) = \beta^2 \Delta_2^2(0\text{-}1)$) (11) the inverse vibrational lifetime T_{LT}^{-1} and (12) the inverse homogeneous lifetime T_{Hom}^{-1}. The calibration mismatch error $\delta\omega_1$ effectively acts like a zero-order frequency shift along the pump axis which accounts for a finite error in the independent calibrations of the pump and probe axes. Though calibration errors are not present in the simulated data, we still treat the fitting as we would with experimental data, where this parameter may be necessary. $\delta\omega_1$ is also useful for fitting to phase distorted data, as shown SI section G. The Kubo amplitude scaling factor β may partially account for situations where dephasing does not scale harmonically between the 0-1 and 1-2 transitions. We tend to find that fitting to inverse lifetimes and squared amplitudes, i.e. T_{Hom}^{-1} and Δ^2, is more stable than T_{Hom} and Δ, which is not surprising given that T_{Hom}^{-1} and Δ^2 appear linearly in exponential arguments of the response function model. At the end of every iteration the program checks to ensure that the next movement $\mathbf{p} + \mathbf{\Delta p}$ is within user-defined boundaries. This routine could be modified to ensure that $T_{hom}^{-1} > \frac{1}{2}T_{LT}^{-1} + \frac{1}{3}T_{or}^{-1}$, which is a physical requirement.[13] For the present cases of isotropic polarization, however, T_{or} is not known *a priori*, so we settle for $T_{hom}^{-1} > \frac{1}{2}T_{LT}^{-1}$ in our boundary checks. Usually, $\frac{1}{3}T_{or}^{-1}$ doesn't contribute much to T_{hom}^{-1}, so we think this is a reasonable approximation for convenience, though model fitting to polarization dependent spectra to fit T_{or} has been done.[17]

For comparison, we also analyze the simulated data using the center-line-slope (CLS) method.[12, 13] One feature of 2D Kubo lineshapes is that they are asymmetric in frequency and the asymmetry is, itself, frequency dependent (see Figure S1). This effect leads to inaccurate center-line measurements when fitting the slices with a symmetric function. Therefore, we fit an asymmetric Lorentzian (Lorentzian + linear term + offset) to slices along the probe axis and then measure the peak of the asymmetric fitting function by numerically finding the extremum of the interpolated function to obtain the center-line points. This method provides a more accurate measure of the true centerline and perfectly agrees with Falvo's analytical expression for calculating the CLS decay by double integration of the response function[48] when tested on our simulated spectra. Fitting a biexponential model to the CLS decay yields the Kubo time constants. For this analysis, we include all waiting time points, including $T_W = 0$. We measure homogeneous dephasing and Kubo amplitudes by fitting to the upper 80% of the linear absorption spectrum holding the Kubo decay times constant. Many baseline distortions are known to occur in FTIR measurements including interference fringes, atmospheric absorption, and scattering from scratched windows, particles and aggregates.[49] Any subtle variation between background and sample measurements, such as path length, index of refraction, concentration and location of particles or aggregates, window scratches, temperature, pressure, atmosphere, and other conditions, can lead to residual distortions in background-subtracted FTIR spectra. These distortions, in addition to various forms of noise, are nontrivial to model so we do not include them in our simulation. In fact, to consider the best-case scenario for the CLS, we fit to a noiseless simulation of the linear absorption spectrum. Nevertheless, we note that in real samples, these contributions would all lead to even greater uncertainty for the lineshape parameters from the CLS method.
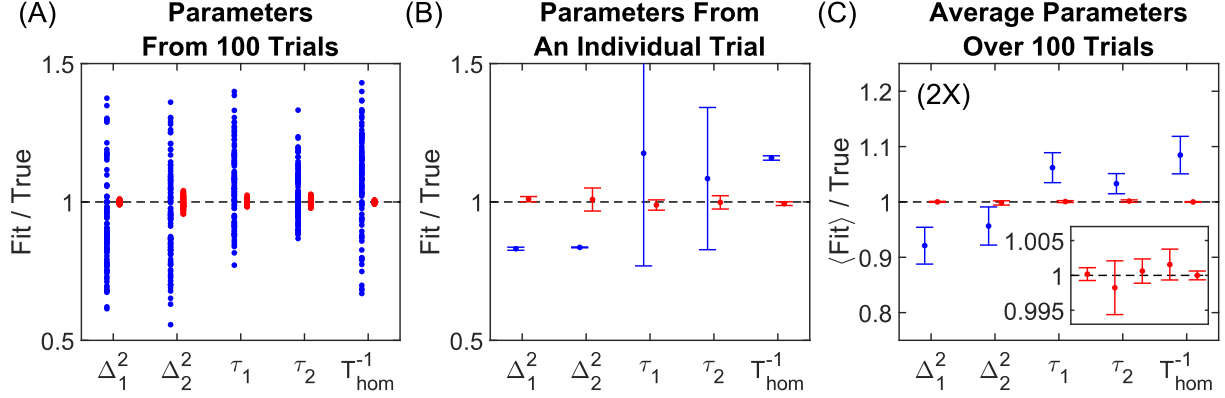
Figure 1. Dephasing parameters by CLS method (blue) and model fitting (red) of simulated 2D IR waiting time series. (A) Fit parameters from all 100 trials. (B) Fit parameters obtained from an individual trial (chosen at random) with 95% confidence intervals estimated from covariance of fit. (C) Average of fit parameters over all 100 trials with 95% confidence interval calculated from standard error of the mean, with inset showing a zoom-in of the true-value line.

Figure 1 shows side-by-side comparisons of the lineshape parameters from both the CLS method (blue) and model fitting (red). For each trial, the code generates a new sampling of random noise for the spectra and chooses a random starting point for fitting **p**. Each panel shows, from left to right, the lineshape parameters: the squared amplitude of the first and second Kubo components (Δ_1^2 and Δ_2^2), the correlation time of the first and second Kubo components (τ_1 and τ_2), and the inverse homogeneous lifetime (T_{Hom}^{-1}). Panel (A) shows the results of all 100 trials plotted as the ratio of fit to true value. Video S1 shows fits to each of the CLS decays for all 100 trials.

Panel (B) shows the results for an individual trial (selected at random) with 95% confidence intervals derived from the covariance of the fit in Eq. 7. Corresponding plots of (B) for every single trial (Video S2) confirm that, for the model fitting approach, the true values fall within the confidence intervals for ~95% of trials. In contrast, the CLS method grossly underestimates the confidence intervals for the Kubo amplitudes and homogeneous dephasing. In calculating these confidence intervals for the CLS method, it was assumed that the Kubo time constants were known with absolute certainty, which is currently the standard treatment. In reality, the Kubo time constants obtained from CLS fits are uncertain, with reported errors ranging from 5-50%.[45, 46, 50, 51] This additional uncertainty must be accounted for when fitting the linear absorption spectrum if we hope to achieve accurate uncertainties for homogeneous dephasing and Kubo amplitudes. Accounting for this effect involves propagating the uncertainty of the Kubo time constants through the standard variance-covariance matrix of the linear absorption fit to obtain the *modified* variance-covariance matrix (Eq. S14), which we derive in supporting information section E. To the best of our knowledge, CLS error bars have never been reported using Eq. S14.

Panel (C) shows the average values and 95% confidence intervals over 100 trials where model fitting improves precision over the CLS method by $8 \times$ to $15 \times$ for Kubo time constants and $8 \times$ to $50 \times$ for Kubo amplitudes and homogeneous dephasing. That true values fall within the intervals in (C) suggests the model fitting is highly accurate and reliable when provided an
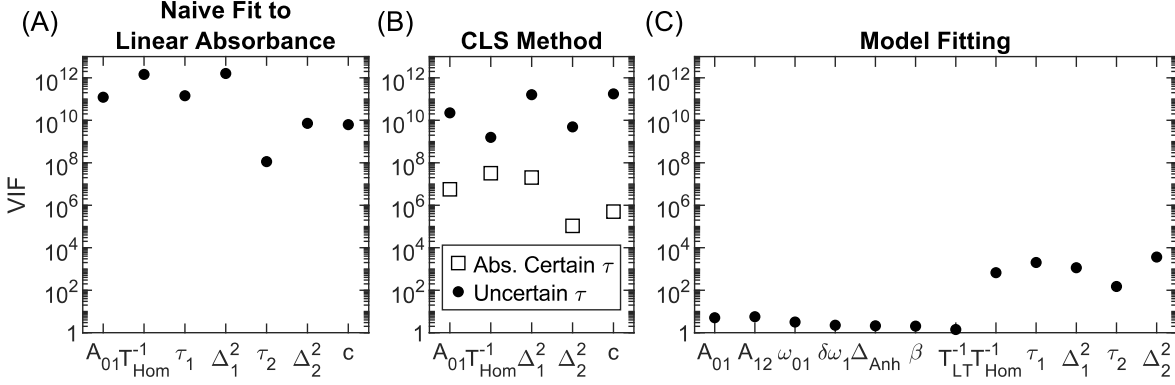
11

Figure 2. The variance inflation factor (VIF) is a measure of multicollinearity (or ill-conditioning) in least-squares regression. Plots of VIFs for (A) naively fitting all dephasing parameters (with scaling and offset) to a linear absorption spectrum, (B) the CLS method and (C) model fitting. Larger VIFs reflect higher multicollinearity.

appropriate model for the data. On the other hand, the CLS method is inaccurate by 5-10% on average.

Propagation of error difficulties aside, fitting dephasing parameters to the linear absorption spectrum is severely ill-conditioned even with the constraints afforded by the CLS method. Figure 2 shows VIFs (see Material and Methods section D) for three different scenarios. Panel (A) is for naively fitting to the upper 80% of the linear absorbance spectrum with a linear response model that includes an amplitude (A_{01}), all five dephasing parameters ($T_{Hom}^{-1}, \tau_1, \Delta_1^2, \tau_2, \Delta_2^2$), and a constant offset (c), which is a scenario that is well known to yield poorly constrained fitting parameters. The resulting VIFs range between 10^8 and 10^{12}. As an example, consider the result of Δ_1^2 with a VIF of 10^{12} which is the inflation that is expected in the variance of Δ_1^2 because one, or more, covariances exist between Δ_1^2 and other fitting variables. If the other six parameters were held constant while fitting Δ_1^2, then the uncertainty of Δ_1^2 should decrease by a factor of 1,000,000 (i.e. $\sqrt{10^{12}}$). In this scenario, the uncertainty of every parameter is so large that any tiny perturbation in the noise causes massive fluctuations in values of the fitting parameters. These results clearly demonstrate that naively fitting all dephasing parameters to the linear absorbance spectrum is an ill-conditioned problem, consistent with prior expectations.[40, 52]

Panel (B) shows results for fitting to the upper 80% of the linear absorbance spectrum using constraints on the time constants similar to what is done in the CLS method. Here we consider the cases where the Kubo time constants are known with absolute certainty (squares) or 10% uncertainty (circles). For the more realistic case of uncertain time constants, VIFs range between 10^9 and ~10^{11}, which imply roughly ~100,000× inflation in the uncertainties of Kubo amplitudes and homogeneous dephasing, give or take an order of magnitude. This result explains the large variance seen in panel (A) of Figure 1, and the necessity of simulating data with a 600:1 SNR in the 2D IR spectra to obtain reliable results for the CLS method. VIFs differ by four to five orders of magnitude between absolutely certain and uncertain Kubo time constants, which implies error bars on Kubo amplitudes and homogeneous dephasing are underestimated by two orders of

12

magnitude. Indeed, this result is observed for the CLS method in Figure 1(B) and Video S2. Finally, panel (C) shows VIFs for model fitting to the 2D IR waiting-time series. For non-dephasing parameters, including the vibrational lifetime, all VIFs are less than 10, meaning multicollinearity is negligible for these parameters. On the other hand, VIFs range between 100 and 3,000 for the remaining dephasing parameters, implying inflated uncertainties between 30× to 50×. Though significant, the inflation of uncertainties in Kubo amplitudes and homogeneous dephasing remains two to four orders of magnitude smaller than for the CLS method.

In addition to comparing model fitting to the CLS method, we are also interested in evaluating the robustness of the model fitting approach to variations in the chosen model, which may not always be known *a priori*. Figure 3 shows plots of $C(\mathbf{p})$ and $|\widetilde{\nabla C}|$ versus fitting iteration for three scenarios in which the fitting model $\mathbf{M}(\mathbf{p})$ has one less (left), the same number (middle), and one more (right) Kubo component(s) than are actually present in the simulated data $\mathbf{D}$. We start with analysis of the middle panels (C) and (D), which correspond to the fitting results shown in Figure 1. Plots of $C(\mathbf{p})$ in panel (C) show that $C(\mathbf{p})$ is useful for broadly comparing the quality of fit for a given iteration of $\mathbf{p}$ but is not particularly useful for understanding convergence because progress near the global minimum is difficult to evaluate on a relative scale. On the other hand, $|\widetilde{\nabla C}|$ shows simple and reproducible behavior that clearly indicates convergence. The variation of $|\widetilde{\nabla C}|$ with iteration shows three phases. First, when $\mathbf{p}$ is far away from a minimum of the cost function, the gradient is $\sim 10^{-1}$. Next, $\mathbf{p}$ reaches the
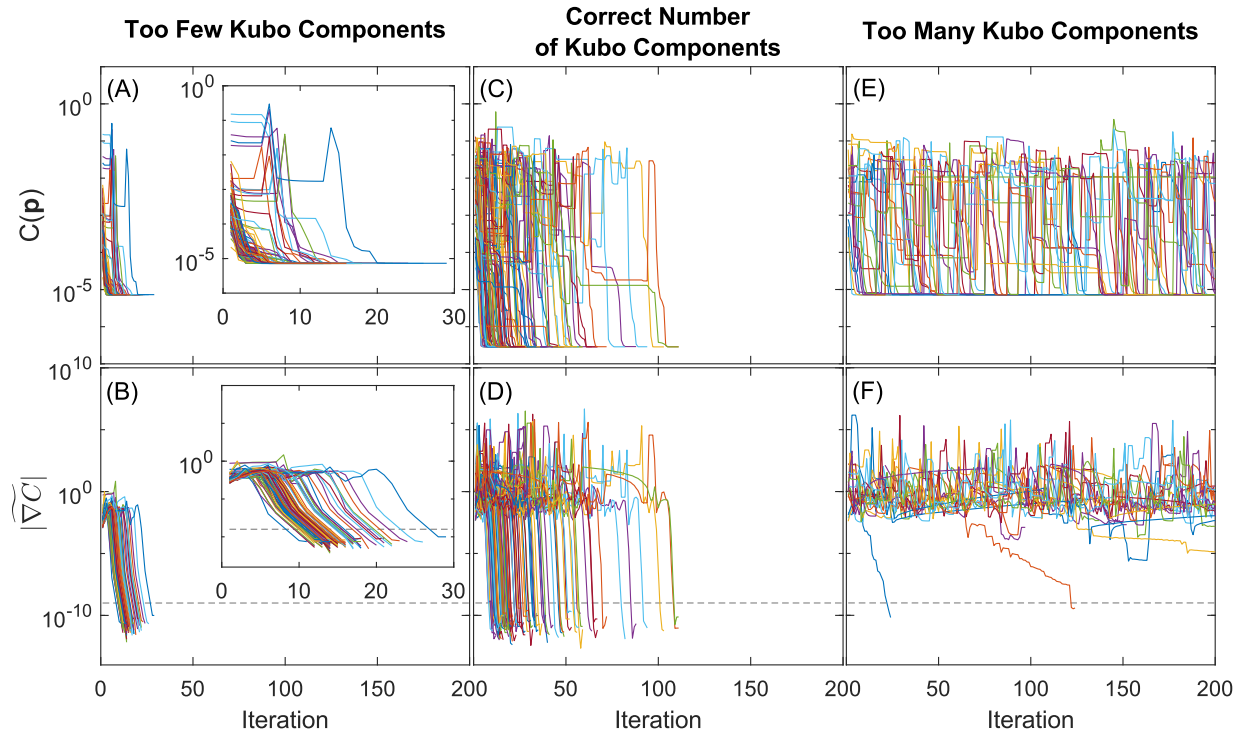

Figure 3. Plots of the cost function and SIGN for modeling with too few (A, B), the correct number (C, D), and too many (E, F) Kubo components. Each trial corresponds to a random sampling of noise and random starting point for $\mathbf{p}$. Dashed line is associated with stopping criterion.

13

neighborhood of a minimum where the second order approximation of $C(\mathbf{p})$ in Eq. 3 is highly accurate, and hence, $|\widetilde{\nabla C}|$ rapidly descends many orders of magnitude in as little as 2-3 iterations. Finally, $\mathbf{p}$ reaches the minimum of $C(\mathbf{p})$ but $|\widetilde{\nabla C}|$ approaches an asymptote. This asymptote should be essentially invariant at the global minimum because $|\widetilde{\nabla C}|$ is, effectively, the relative roundoff error in $C(\mathbf{p})$, and hence, is determined by accumulated roundoff errors in Eq. 8 on top of machine precision ($\sim 2 \times 10^{-16}$). In a few cases, $|\widetilde{\nabla C}|$ descends halfway and stalls. The algorithm responds to stalling by restarting from a new randomly generated $\mathbf{p}$ within the user-supplied boundaries corresponding to the sudden increases seen in both plots (C) and (D). The dashed line in (C) is a user defined limit associated with the stopping criteria (section C of Materials and Methods). Video S3 shows fitting trajectories for every trial.

Panels (A) and (B) show plots of $C(\mathbf{p})$ and $|\widetilde{\nabla C}|$ for 100 trials of fitting a one-Kubo model to data simulated by a two-Kubo lineshape. Compared to (C) and (D), fitting is far less susceptible to stalling and converges significantly faster, which suggests $C(\mathbf{p})$ is more convex over a wider range of $\mathbf{p}$. Past a point of $|\widetilde{\nabla C}| \approx 10^{-7}$, the second order approximation of $C(\mathbf{p})$ is less effective and converges at a moderately slower rate. Fitting trajectories in Video S4 show $\mathbf{p}$ consistently reaches the same global minimum for all 100 trials.

Panels (E) and (F) show plots of $C(\mathbf{p})$ and $|\widetilde{\nabla C}|$ for 20 trials of fitting a two-Kubo model to data simulated by a one-Kubo lineshape. The many jumps in both plots are evidence of frequent stalling and random restarts, which is opposite of the behavior seen in panels (A) and (B). We ran this experiment for only 20 trials due to the large number of iterations, the resulting crowding in the plots, and the larger memory requirement for the video.

Fitting trajectories in Video S5 show many examples of stalling and abnormal convergence corresponding to panels (E) and (F) in Figure 3. Stalling occurs when a single fitting iteration, ending with $|\Delta\mathbf{p}| \approx 0$, starts an insidious cycle where $\mathbf{p}$ doesn't change and so the algorithm is doomed to repeat itself barring some intervention. This is distinct from local or global minima because stalling also requires $|\widetilde{\nabla C}| > 10^{-9}$, which implies $\mathbf{p}$ is not a minimum of $C(\mathbf{p})$. Two scenarios for which stalling may occur are (1) $\Delta\mathbf{p}$ is accurate in direction but $|\Delta\mathbf{p}|$ is limited by a boundary condition which causes $|\Delta\mathbf{p}| \approx 0$ or (2) multicollinearity among fitting parameters leads to a nearly singular $\nabla\nabla C$ which results in an erroneous vector direction of $\Delta\mathbf{p}$ (Eq. 6) and hence, $\Delta\mathbf{p}$ is unable to reduce the cost function and so $|\Delta\mathbf{p}| \approx 0$. Figure 4 illustrates three such examples of stalling. For each example, the top panel (A), (C), or (E) shows fitting trajectories of Kubo amplitudes (y-axis) versus time constants (x-axis) where the dashed line reticles mark the "true" value from the simulation input, and the bottom panel (B), (D), or (F) shows the plot of SIGN versus iteration for the fit. The first example in the left panel of Figure 4 shows the case of boundary stalling, in which the direction $\Delta\mathbf{p}$ is accurate for further minimizing the cost function, but the location $\mathbf{p} + \Delta\mathbf{p}$ is outside of the user-defined boundary for at least one of the parameters (in this case $\tau_c \leq 10$). Convergence ($|\widetilde{\nabla C}| \ll 1$) is impossible in the case of regular boundary stalling and the program readily detects the constant value of $|\widetilde{\nabla C}|$ and resolves the stall with a random restart.
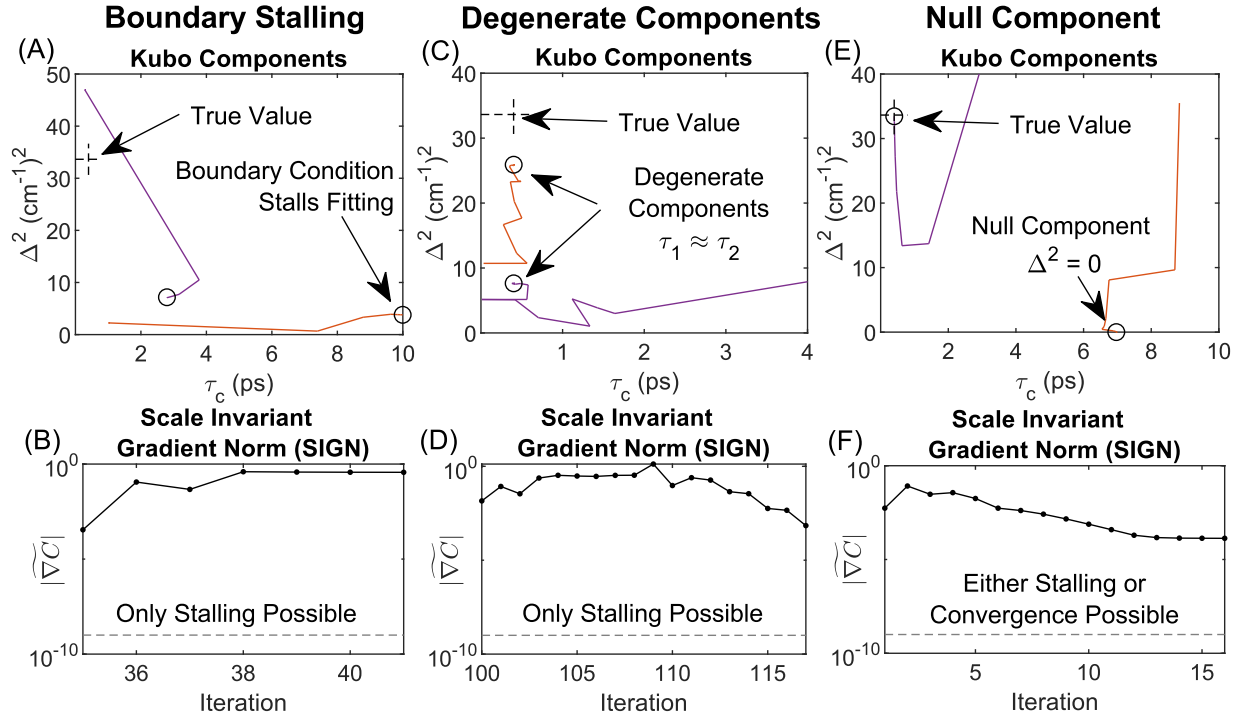
Figure 4. Select examples of stalling or unusual convergence. Plots (A) and (B) show an example of boundary stalling. Plots (C) and (D) show an example of stalling due to multicollinearity between degenerate Kubo components. Plots (E) and (F) show an example in which one component approaches the true value and the other approaches null (i.e. $\Delta^2 = 0$), where either stalling or convergence may occur.

The second example in the center panel shows the formation of a degenerate pair of Kubo components (i.e. having the same time constant). Plots of VIFs shown in Videos S3 and S5 are perfect examples of multicollinearity when a degenerate pair align. When a degenerate pair align ($\tau_1 \approx \tau_2$), the FFCF reduces to $\Delta_1^2 \exp(-t/\tau_1) + \Delta_2^2 \exp(-t/\tau_2) \approx (\Delta_1^2 + \Delta_2^2) \exp(-t/\tau_1)$ and hence, the amplitudes Δ_1^2 and Δ_2^2 are linearly dependent and therefore indistinguishable. As shown in many examples of Video S3 (e.g. trial 5, iteration 26) and Video S5 (e.g. trial 3, iteration 111) the VIFs and condition number always explode to infinity when pairs align. In every case, the program detects the blow up and reacts by random restart.

The third example in the rightmost column shows a null Kubo component (i.e. $\Delta^2 = 0$), which is a special case of boundary stalling. It is tempting to think convergence with $\Delta^2 = 0$ should be possible, and examples of this are in fact observed (e.g. Video S5, trial 1, iteration 24), but there are several cases in which stalling occurs instead (e.g. Video S5, trial 5, iteration 26). Numerical analysis (not shown) reveals that stalling here is caused by inflated estimates of $|\widetilde{\nabla C}|$ stemming from near unity covariance(s) with the null component τ_c, which is a form of multicollinearity. This is not surprising given the uncertainty of τ_c is infinite for a true null component. In any case, we did not identify a scenario in which the program converged to an inaccurate fit after accounting for degenerate pairs, which implies that local minima are extremely rare for a three-level system assuming accurately phased spectra.
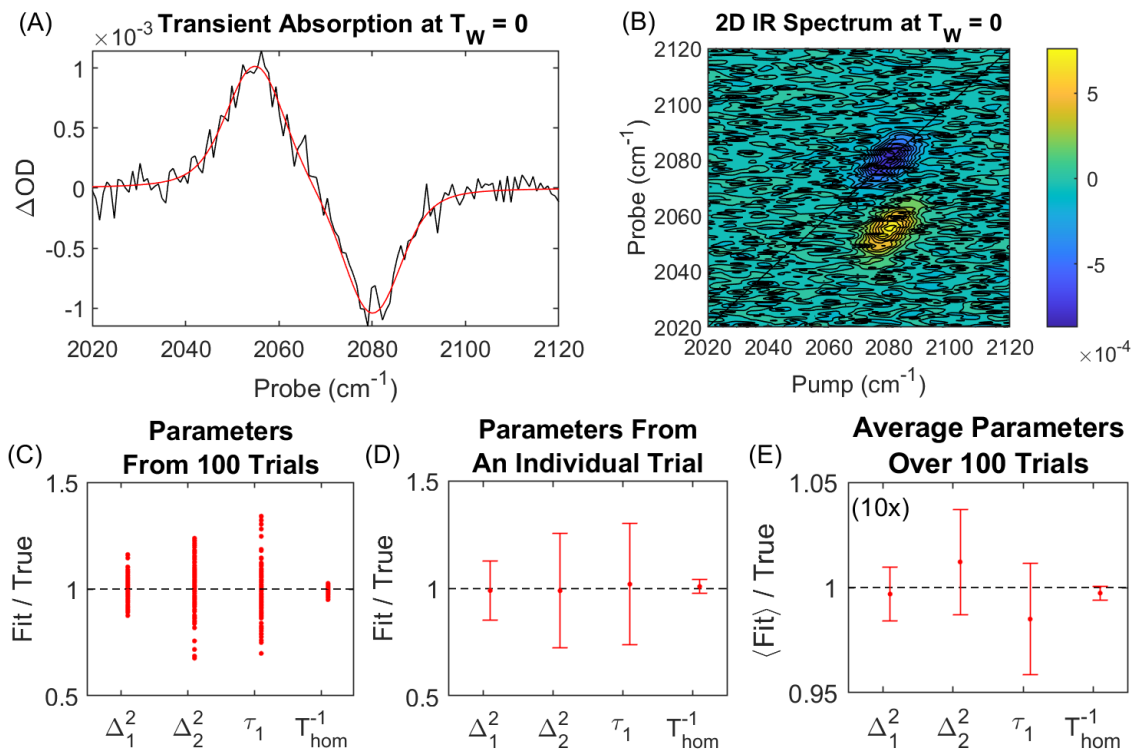
Figure 5. Model Fitting to low SNR 2D IR waiting-time series. Simulation is representative of a cyanylated cysteine residue in the protein Calmodulin, where Δ_2^2 is static relative to the vibrational lifetime. Examples of (A) transient absorption and (B) 2D spectrum with SNR of 10:1. (C) Fit parameters from 100 trials. (D) Fit parameters obtained from an individual trial with 95% confidence intervals estimated from covariance of fit. (E) Average of fit parameters over all 100 trials with 95% confidence interval calculated from standard error of the mean.

Figure 5 shows the results for fitting to data with a much lower SNR of 10:1. Simulation parameters are representative of a cyanylated cysteine residue in the Calmodulin protein.[53] The second Kubo component is treated as static on the timescale of the waiting-time measurements by holding τ_2 constant at 1 ns. The transient absorption spectrum in (A) and the 2D IR spectrum in (B) at $T_W = 0$ ps illustrate just how modest the quality of the raw data is at 10:1 SNR. Dephasing parameters in (C) from model fitting for all 100 trials show a distribution that rarely exceeds 10-25% of the true values. Any analysis based on the CLS method would be hopeless at this 10:1 SNR. An example of dephasing parameters for an individual trial in (D) shows that the 95% confidence intervals accurately reflect the variance seen in (C), which shows that the calculated uncertainties are reliable for low SNR data. Indeed, each of the individual results for the 100 trials in panel (D) of Video S6 validates the 95% confidence intervals. Panel (E) shows the average over all 100 trials with corresponding 95% confidence intervals calculated from the standard error of the mean confirming the accuracy of model fitting for this example. Fitting trajectories seen in Video S7 are similar to the case of MeSCN in H_2O previously shown in Figure 1 and Video S3 above. Some occasions of stalling do occur which the program readily resolves. As expected, there is a decrease in the precision of the modeling results due to the increase in the noise.
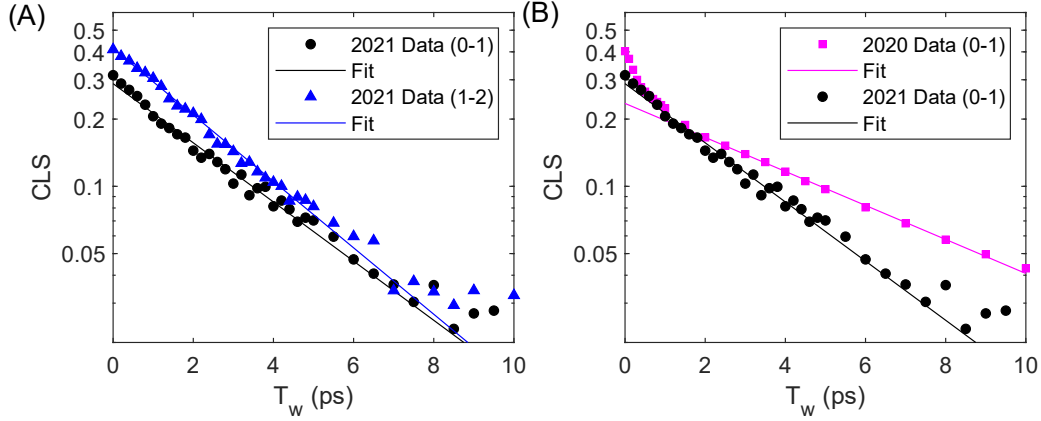
16

Figure 6. CLS of MeSCN in DMSO. (A) Comparison between 2020 data (magenta squares) and 2021 data (black circles). 2021 data is background subtracted prior to CLS analysis while 2020 data is not. (B) Comparison between 0-1 (black circles) and 1-2 (blue triangles) CLS for 2021 data.

B. Fitting Experimental Data

For MeSCN in DMSO, we have measured and analyzed two independent data sets. One comes from our previous publication on edge-pixel referencing, and we refer to these as the 2020 data.[27] We have also collected a new set of data that we will identify as the 2021 data. The original motivation for collecting the 2020 data was to compare data processed by two different referencing schemes. We were, therefore, not motivated to account for the background solvent response. On the other hand, in the 2021 data the pump pulse is lower in energy by ~8× compared to the 2020 measurements, and we have subtracted the solvent background for this measurement. The other notable differences between the data sets are the concentration (2020 is 200 mM, 2021 is 400 mM), the pathlength (2020 is 100 μm, 2021 is 50 μm), and that the 2020 data have significantly higher SNR due to more averaging and larger signal strength because of the higher pump energy. Figure 6 (A) shows the results of the CLS analysis of the two data sets. Having accounted for the background response, the CLS in the 2021 data set exhibits a single exponential decay, in contrast to what is seen in the 2020 data. In addition, the CLS decay times differ by 60% between data sets (5.8 ps vs. 3.3 ps) even if we only fit the long-time scale component for the 2020 data.

Figure 6(B) shows the comparison of the 0-1 and 1-2 CLS decays for the 2021 data set. The two transitions have different initial CLS values (0.3 for the 0-1 and 0.4 for the 1-2) and the CLS decay of the 1-2 transition appears to be marginally faster than the 0-1 peak. We therefore model spectral diffusion of the 0-1 transition as $\langle \delta\omega_{01}(t)\delta\omega_{01}(0)\rangle = \delta(t)/T_{hom} + \Delta^2 \exp(-t/\tau)$ and the 1-2 transition as $\langle \delta\omega_{12}(t)\delta\omega_{12}(0)\rangle = \delta(t)/T_{hom} + \beta^2\Delta^2 \exp(-t/\tau)$ where β is a unitless scaling factor that accounts for the larger 1-2 CLS amplitude seen in Figure 6(B). We also account for the effect that the shorter 1-2 vibrational lifetime has on the 1-2 dephasing during the second coherence time.[40, 44]

Figure 7 shows plots of (A) the cost function and (B) $|\widetilde{VC}|$ for model fitting to the 2020 data. A first attempt to model the data with a two-Kubo lineshape (blue) does not converge and clearly

17

resembles the results in Figure 4(E) and (F) for a model with too many Kubo components. A second attempt to model the data using a one-Kubo model (orange) immediately converges, implying the data are best modeled by a one-Kubo lineshape. Table 1 shows results for all ten fitting variables. Our earlier analysis of the 2020 data reported a vibrational lifetime of 75 ± 4 ps, which agrees with T_{LT}^{-1} reported by model fitting in Table 1 (inverting to 72.5 ± 0.3 ps). On the other hand, our earlier report of 5.8 ± 0.3 ps for the CLS time constant is 60% larger than the Kubo time constant reported by model fitting of the same data, 3.57 ± 0.04 ps. Importantly, however, the model fitting value of the 2020 data agrees well with the CLS time constant of 3.3 ± 0.3 ps for the 2021 data and the model fitting Kubo time constant of 3.28 ± 0.02 ps for the 2021 data.

Figure 8 shows 2D IR spectra of 2020 data in panels (A) and (D) for waiting times of 0.4 ps and 50 ps, and corresponding best-fit model spectra in panels (B) and (E). Qualitatively, the model appears consistent with the data in terms of shape and scale. A closer look at the residuals in panels (C) and (F), however, reveals the presence of a structured response at both waiting times. This residual response is roughly 10% of the amplitude of the total signal and is also present at similar magnitude in the 2021 data.

The waiting-time dependent residuals differ in shape between the 2020 and 2021 data around the 0-1 peak (Videos S8 and S9), but for frequencies less than 2035 cm^{-1} are similar around the diagonal, particularly at early waiting times. Differences around the 0-1 peak likely cause the discrepancies in the CLS decays shown in Figure 6A. To show this more clearly, we overlay the centerlines of the experimental spectra in the left and right panels of Videos S8 and S9. Close inspection of the right panel in Video S8 reveals

#	Parameter	Fitted Value
1	A_{01}	16.56 ± 0.04 arb. unit
2	A_{12}	18.73 ± 0.04 arb. unit
3	ω_{01}	2153.319 ± 0.006 cm^{-1}
4	$\delta\omega_1$	0.137 ± 0.007 cm^{-1}
5	Δ_{Anh}	25.483 ± 0.009 cm^{-1}
6	β	1.138 ± 0.003 unitless
7	T_{LT}^{-1}	0.01379 ± 0.00003 ps^{-1}
8	T_{Hom}^{-1}	0.229 ± 0.003 ps^{-1}
9	τ	3.57 ± 0.04 ps
10	Δ^2	14.8 ± 0.1 cm^{-2}

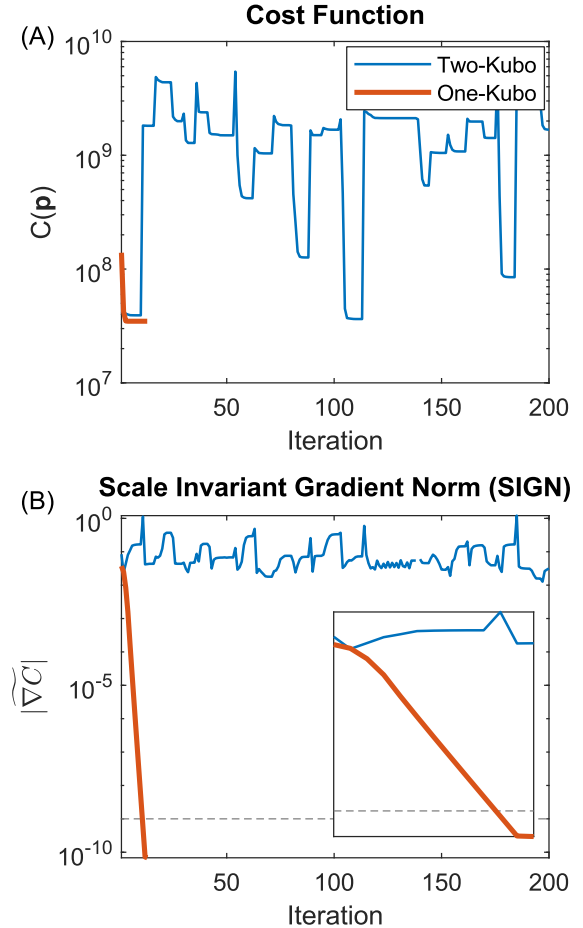Table 1. Fitting parameters for 2020 data.

Figure 7. Plots of (A) cost function and (B) scale invariant gradient norm for one (orange) and two (blue) Kubo component models fitted to 2020 data.
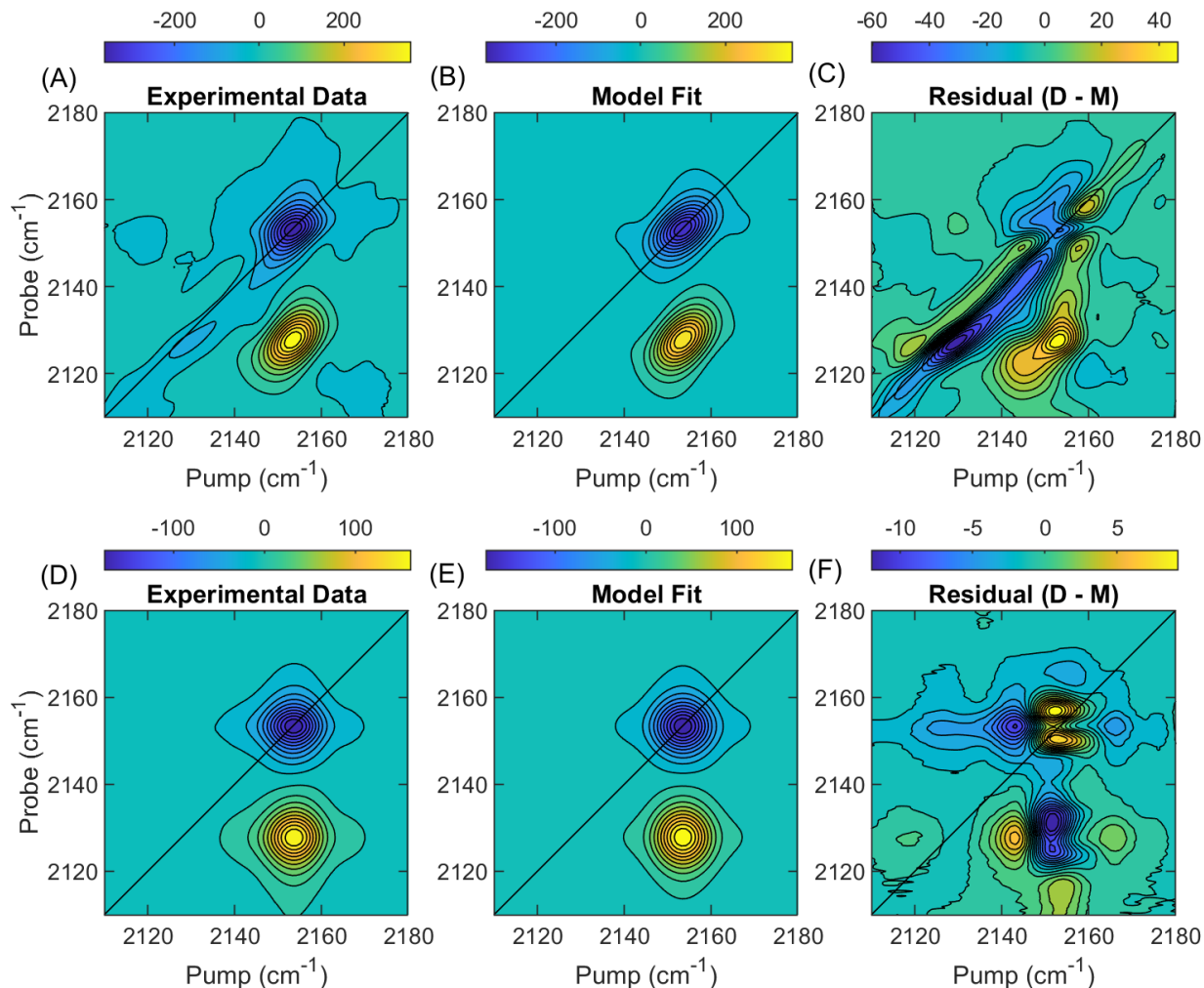
18

Figure 8. Plots of (A and D) 2020 data, (B and E) model fit result and (C and F) residual for T_w= 0.4 ps and 50 ps, respectively. Residual in (C) shows a structured response which is unaccounted for by the model.

the structured feature responsible for the slower CLS decay in the 2020 data which is not present in the 2021 data (Video S9). We must caution against over interpreting the residual here. One should not conclude that the residual reflects the "unmodeled" part of the lineshape where the model $\mathbf{M}(\mathbf{p})$ and Jacobian $\mathbf{J}$ are nonzero. The best fit model is influenced by both "modeled" and "unmodeled" parts of the experimental lineshape such that $C(\mathbf{p})$ is minimized. On the other hand, in the region below 2135cm^{-1} along the diagonal, the residual does accurately reflect the "unmodeled" lineshape because $\mathbf{M}(\mathbf{p})$ and $\mathbf{J}$ are virtually zero and therefore the residual $\mathbf{r}$ in this region cannot influence the model fit by Eq. 4.

We note three interesting features of this lower frequency response at early waiting time. First, as shown in Videos S8 and S9, the phase of the signal oscillates as a function of waiting time with a ~1.3 ps period (~25 cm^{-1}). Second, the signal initially appears stretched along the diagonal and dephases with a lifetime of 1-2 ps. Third, as shown in both videos, the intensity of the residual response is roughly 10% of the peak 2D IR signal in both 2020 and 2021 data, which
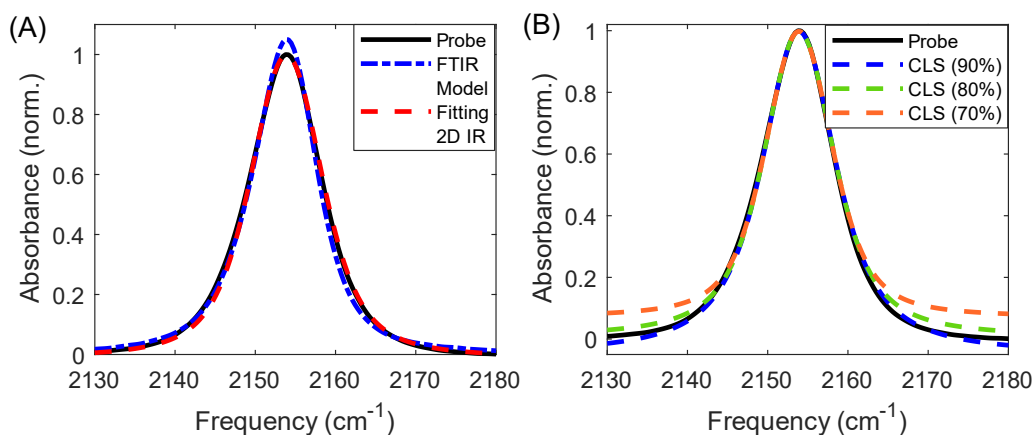
19

Figure 9. (A) Linear absorption spectra as measured by the probe (black, solid), FTIR (blue, dot-dashed), and simulated with parameters obtained by model fitting the 2D IR waiting time series (red, dashed). (B) Results of the CLS method fitting to the upper 90% (blue, dashed), 80% (green, dashed) and 70% (orange, dashed) of the probe spectrum (black, solid) to obtain the Kubo amplitude and homogeneous dephasing.

implies a third-order response. We suspect this may be a resonantly enhanced wave packet of low-frequency Raman modes (e.g. the methyl torsion) anharmonically coupled to the C≡N stretch, similar to what has been seen in a variety of other oscillators.[54-56] At longer waiting times, the residual appears as a vertical peak shift of the 0-1 transition coincident with the vibrational lifetime, which is characteristic of hot ground state absorption.[56-59]

Figure 9(A) shows the linear absorbance spectrum for 2021 data obtained using the probe beam and upconversion spectrometer (black, solid) and using an FTIR (blue, dashed). We subtracted DMSO backgrounds in both spectra, offset and scale the FTIR spectrum to best match the probe spectrum via linear least-squares, but the FTIR spectrum remains notably narrower than the probe spectrum. We attribute this to different instrument response functions between the two measurements. For example, the FTIR spectrum is influence by factors such as apodization, scan length and vignetting of light along moving optics,[49] while the probe spectrum is influenced by the resolving power of the spectrometer and the bandwidth of the 800 nm pump utilized in upconverting the infrared light prior to the spectrometer.

The third trace in Figure 9A (red, dashed) is the linear response predicted by model fitting to the 2D IR waiting time series (scaled and offset to best match the spectrum measured with the probe beam). The strong match between the predicted response and probe spectrum is independent validation of model fitting. For comparison, Figure 9B shows linear absorption fits to the FTIR spectrum for the CLS method. FTIR lineshapes can be unreliable near the baseline due to the difficulties with imperfect background subtraction, especially for dilute solutions of weak chromophores, and therefore, we show results for fitting to the upper 90%, 80% and 70% of the linear absorbance spectrum. We float the homogeneous lifetime, Kubo amplitude, linear scaling and offset as fitting variables while holding the Kubo correlation time constant at 3.3ps. We see that 80% provides a reasonable balance between quality of fit and distortions as a result of the baseline, which is in line with the recommendation by Kwak and coworkers.[13] It is notable
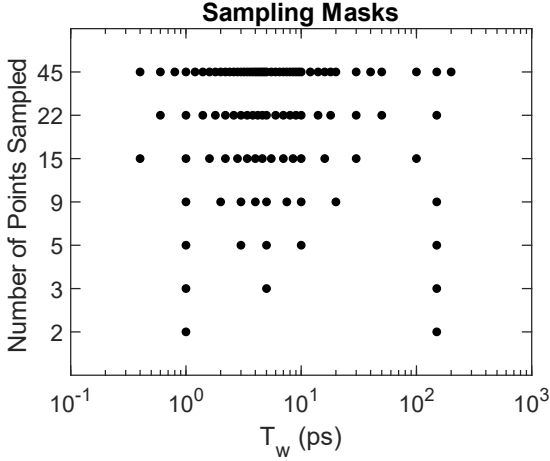
Sampling Masks

Figure 10. Masks used in undersampling the waiting time T_W of MeSCN in DMSO. Each row corresponds to a different sampling mask. Black dots represent the inclusion of a 2D spectrum in fitting for a given waiting time. In all cases, we exclude the first 300 fs of waiting time to avoiding spurious nonresonant and time-ordering signals from pulse overlap.

that the linear absorption predicted by model fitting in Figure 9A is still a higher quality fit than the 90% case for the CLS analysis.

We now examine the efficacy of model fitting of undersampled data with the sampling masks illustrated in Figure 10. There are 47 waiting time spectra in the entire series, but the first two (T_w = 0 and 200 fs), where pump and probe overlap, are susceptible to spurious nonresonant and time-ordering signals. Hence, we omit these spectra from model fitting in each case and "fully sampled" data corresponds to the 45-point mask. We generate masks by keeping every k^{th} point left and right of 1 ps, where k is an integer number. For example, the 15-point mask samples every third point left and right of 1 ps. We base this on the guiding principle that every mask should include a waiting time point ~2× earlier than the shortest process expected to occur in the lineshape and ~2× later than the vibrational lifetime. Knowing *a priori* that the homogeneous lifetime is ~2 ps and vibrational lifetime is ~75 ps, we include points around 1 ps and 150 ps in every mask. For other -SCN systems more generally, one can reasonably assume that the shortest observable process is likely no faster than ~0.8 ps and the vibrational lifetime is between 30-80 ps.

Figure 11 shows plots of $C(\mathbf{p})$ and $|\widetilde{\nabla C}|$ for all sampling masks. As noted earlier, the cost function scales linearly with N_D, which explains

(A)

Cost Function

——	45 T_w points
——	22 T_w points
——	15 T_w points
——	9 T_w points
——	5 T_w points
——	3 T_w points
——	2 T_w points

(B)

Scale Invariant Gradient Norm (SIGN)

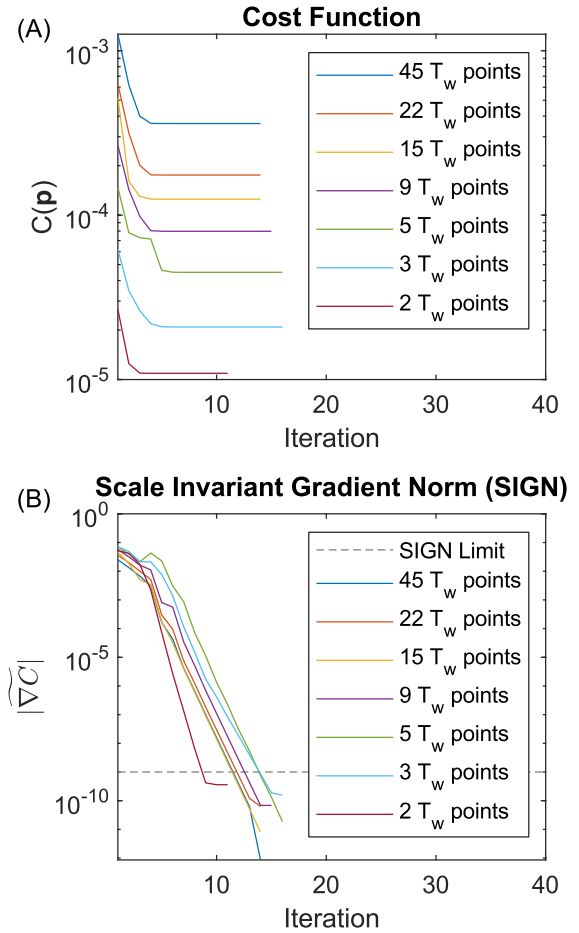|	
- - - -	SIGN Limit
——	45 T_w points
——	22 T_w points
——	15 T_w points
——	9 T_w points
——	5 T_w points
——	3 T_w points
——	2 T_w points

Figure 11. Plots of (A) the cost function C and (B) the scale invariant gradient norm $|\widetilde{\nabla C}|$ versus fitting iteration for series of undersampled waiting time T_w. Dashed line in (B) is associated with stopping criterion.

21

the varying magnitudes of $C(\mathbf{p})$ in Figure 11(A). On the other hand, the scale invariant gradient norm has no dependence on N_D, and hence, $|\widetilde{\nabla C}|$ is similar in scale for every mask.

Figure 12 shows dephasing parameters obtained by the CLS method (left column) and model fitting for every sampling mask (right column). The 60% discrepancy between Kubo time constants in panel (A) is a reflection of the CLS decays in Figure 6(A). In contrast, Kubo time constants obtained by model fitting in (D) differ by just 10% between the 2020 and 2021 data sets. This shows that model fitting is far more consistent and reliable for estimating Kubo time constants than the CLS method. The results in (D) also show that Kubo time constants are consistent across all undersampled versions of 2021 data. It is remarkable that model fitting to just two waiting time spectra yields precision comparable to that of CLS method for 45 waiting time points. In practice, we do not recommend fitting to only two waiting time spectra as we would not expect this to work for multi-Kubo lineshapes.

Kubo amplitudes and homogeneous dephasing obtained by the CLS method shown in (B) and (C) reflect the linear absorption fits in Figure 9(B). Kubo amplitudes and homogeneous dephasing obtained by model fitting in (E) and (F) show consistent values across all undersampled versions of 2021 data. There does appear to be a systematic depression of 15% in the homogeneous dephasing rate from left to right, but this difference is small in comparison to the uncertainty of the CLS method. Finally, results in (F) also show a 40% discrepancy between homogeneous dephasing rates obtained by model fitting between the 2020 and 2021 data sets. Some of this difference may be the result of background subtracting only the 2021 data set, or inaccuracies in inverting the apodization window required for the 2020 data (i.e. Materials and Methods Section E). Nonetheless, this 40% discrepancy is still small in comparison to the uncertainty seen in the CLS method.
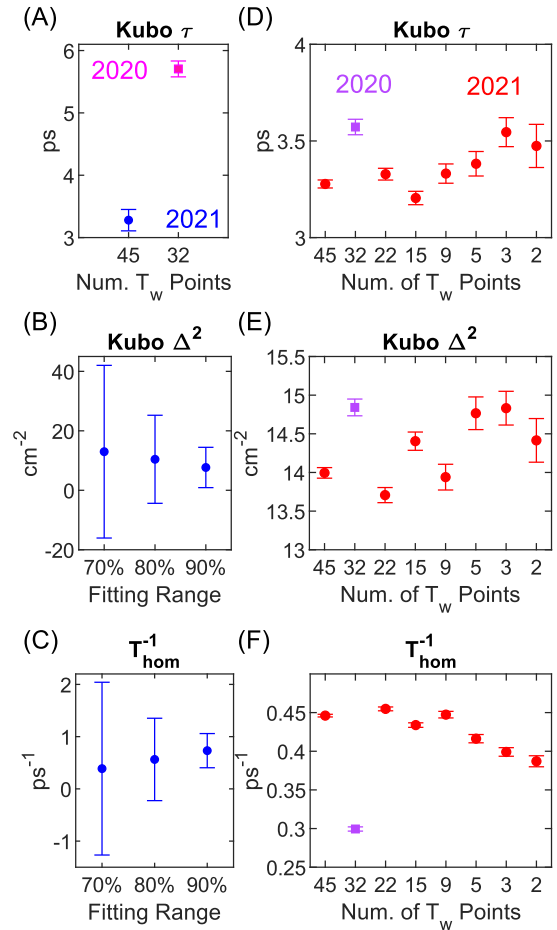

Figure 12. Comparison of dephasing parameters obtained by CLS Method and Modeling Fitting. Plot (A) shows Kubo time constants obtained by CLS method for 2020 and 2021 data. Plots (B) and (C) are Kubo amplitudes and homogeneous dephasing obtained by CLS method for 2021 data. This is plotted for three different fitting ranges of linear absorption (see Figure 9B) to demonstrate how sensitive these parameters are to the linear absorption spectrum. Plots (D), (E) and (F) correspond to model fitting of 2021 data as a function of the number of waiting time points used in fitting. Model fitting to 2020 data also shown for comparison. Error bars are 95% confidence intervals estimated from covariance of fit.

Conventional wisdom says that model fitting of multidimensional spectra is a problem riddled by local minima. In the several hundred trial examples of model fitting shown here, our program encounters and resolves many algorithmic stalls, but not a single local minimum is observed. The distinction is important. It is difficult, and often impossible, to distinguish a local minimum from the global minimum, but stalling is always distinguishable using $|\widetilde{\nabla C}|$. Though we have only shown that local minima are exceptionally rare for a simple three-level system, it is reasonable to believe this should apply to more advanced models including coupled oscillators, assuming reasonable separation between peaks, and more complicated lineshape functions. Therefore, conventional wisdom should be updated: model fitting is far less a problem of local minima as it is of multicollinearity and boundaries, which are manageable.

C. Recommended Practices for Model Fitting

We summarize the following recommendations for model fitting.

1. For faster performance, limit the number of data points. See Materials and Method Section F for suggestions.
2. We strongly discourage superfluous zero padding or apodizing data prior to fitting as these effects will propagate into fitting results. We recommend fitting to data in the original measurement domain. See Materials and Methods section E for more information.
3. When applicable, users should enable a zero-order phase fitting parameter to account for residual phasing errors (see SI section G). The presence of spectrally correlated shot-to-shot noise (a.k.a. local-oscillator noise) likely limits the ability to accurately phase data, particularly with the projection slice theorem. Therefore, we do not recommend model fitting for phase distorted apparatuses without prior removal of shot-to-shot noise.
4. Calibrated referencing schemes greatly improve precision of fitting parameters by removing correlated shot-to-shot noise. For example, we found edge-pixel referencing reduced uncertainties by a factor of 10 over unreferenced data.
5. The relative variance across data (e.g. due to non-uniform averaging across different waiting time spectra) should be accounted for using V_D^{-1}. This is straight-forward using our GUI interface. See Material and Methods section B for more information.
6. Fitting happens simultaneously across all spectra, and therefore, consistency is required among factors that affect the magnitude of the nonlinear signal throughout the entire experiment (e.g. constant pump power, constant intensity ratio between probe and local oscillator, consistent normalization factors if non-uniformly averaging along waiting time spectra).
7. For best performance, tune the initial parameters to reasonably match the model and data prior to fitting. Our GUI interface provides real-time update of the lineshape model for comparison with the data across several different plots to greatly help with this process.
8. For best performance, boundary conditions should strike a reasonable balance between narrow enough to avoid too many random restarts and wide enough to avoid boundary stalling. Again, our GUI interface is very helpful here.

9. Start by fitting a one-Kubo model and then add more components. For best performance with two or more components, avoid degeneracy by restricting the boundary conditions of each time component to a different time scale (e.g. 0.1 – 2.0 ps and 1.0 – 10 ps). Use the time constant from the one-Kubo component fit to estimate the dividing line between the two time scales with comfortable overlap to avoid over constraining. With the exception of null components, convergence will not occur for models having more Kubo components than truly present in the data. Note that it is practically impossible to resolve a Kubo correlation time much longer than the vibrational lifetime due to low SNR at longer T_w, which is similarly true for the CLS method.
10. When exploring different models, VIFs are a useful tool for monitoring multicollinearity. Plots of VIFs are readily generated in our GUI app.

Conclusion

We introduce a scale invariant gradient norm (SIGN) capable of identifying, and distinguishing between, algorithmic stalling and convergence at a local or global minimum. Our model fitting algorithm accurately estimates all lineshape parameters with superior precision and accuracy compared to the CLS method. We show how to infer when a model has too many, or too few, Kubo components for a data set based on the behavior of the SIGN. Interestingly, we find no evidence of local minima when fitting to a multi-Kubo lineshape of a three-level system.

Though analysis of simulated spectra suggests the CLS method is reliable in retrieving Kubo time constants, we have shown an experimental example in which the CLS time constants differ by 60% between independent measurements of the same system. In contrast, Kubo time constants obtained by model fitting only differ by 10%, which suggests model fitting is a far more reliable and consistent means of measuring spectral diffusion over the CLS method. Furthermore, we revealed a fundamental oversight in propagation of error with the CLS method, which led us to show that error bars for Kubo amplitudes and homogeneous dephasing obtained by fitting the linear absorbance spectrum are unreliable. In contrast, model fitting yields reliable error bars over a wide range of scenarios with upwards of 50× better precision than the CLS method.

While the scope of this study is limited to the isotropic response of a simple three-level system, we expect model fitting to the anisotropic response should work equally as well. We plan to explore this in a follow up study in addition to more complicated lineshape models, such as coupled oscillators, overlapping ensembles, and underdamped oscillatory FFCFs.

Acknowledgements

References

1. Kubo, R., A stochastic theory of line shape. *Stochastic processes in chemical physics* **1969,** *15*, 101-127.
2. Demirdöven, N.; Khalil, M.; Tokmakoff, A., Correlated vibrational dynamics revealed by two-dimensional infrared spectroscopy. *Physical review letters* **2002,** *89* (23), 237401.
3. Hybl, J. D.; Yu, A.; Farrow, D. A.; Jonas, D. M., Polar solvation dynamics in the femtosecond evolution of two-dimensional Fourier transform spectra. *The Journal of Physical Chemistry A* **2002,** *106* (34), 7651-7654.
4. Eaves, J.; Loparo, J.; Fecko, C. J.; Roberts, S.; Tokmakoff, A.; Geissler, P., Hydrogen bonds in liquid water are broken only fleetingly. *Proceedings of the National Academy of Sciences* **2005,** *102* (37), 13019-13022.
5. Asbury, J. B.; Steinel, T.; Stromberg, C.; Corcelli, S.; Lawrence, C.; Skinner, J.; Fayer, M., Water dynamics: Vibrational echo correlation spectroscopy and comparison to molecular dynamics simulations. *The Journal of Physical Chemistry A* **2004,** *108* (7), 1107-1119.
6. Lazonder, K.; Pshenichnikov, M. S.; Wiersma, D. A., Easy interpretation of optical two-dimensional correlation spectra. *Optics letters* **2006,** *31* (22), 3354-3356.
7. Lazonder, K.; Pshenichnikov, M. S.; Wiersma, D. A., Ultrafast Phenomena XV. 2007; Vol. 88.
8. Hamm, P., Three-dimensional-IR spectroscopy: Beyond the two-point frequency fluctuation correlation function. *The Journal of chemical physics* **2006,** *124* (12), 124506.
9. Loparo, J. J.; Roberts, S. T.; Tokmakoff, A., Multidimensional infrared spectroscopy of water. II. Hydrogen bond switching dynamics. *The Journal of chemical physics* **2006,** *125* (19), 194522.
10. Roberts, S. T.; Loparo, J. J.; Tokmakoff, A., Characterization of spectral diffusion from two-dimensional line shapes. *The Journal of chemical physics* **2006,** *125* (8), 084502.
11. Finkelstein, I. J.; Ishikawa, H.; Kim, S.; Massari, A. M.; Fayer, M., Substrate binding and protein conformational dynamics measured by 2D-IR vibrational echo spectroscopy. *Proceedings of the National Academy of Sciences* **2007,** *104* (8), 2637-2642.
12. Kwak, K.; Rosenfeld, D. E.; Fayer, M., Taking apart the two-dimensional infrared vibrational echo spectra: More information and elimination of distortions. *The Journal of chemical physics* **2008,** *128* (20), 204505.
13. Kwak, K.; Park, S.; Finkelstein, I. J.; Fayer, M. D., Frequency-frequency correlation functions and apodization in two-dimensional infrared vibrational echo spectroscopy: A new approach. *The Journal of Chemical Physics* **2007,** *127* (12), 124503.
14. Park, S.; Kwak, K.; Fayer, M., Ultrafast 2D-IR vibrational echo spectroscopy: A probe of molecular dynamics. *Laser Physics Letters* **2007,** *4* (10), 704.
15. Candelaresi, M.; Gumiero, A.; Adamczyk, K.; Robb, K.; Bellota-Antón, C.; Sangal, V.; Munnoch, J.; Greetham, G. M.; Towrie, M.; Hoskisson, P. A., A structural and dynamic investigation of the inhibition of catalase by nitric oxide. *Organic & Biomolecular Chemistry* **2013,** *11* (44), 7778-7788.
16. Guo, Q.; Pagano, P.; Li, Y. L.; Kohen, A.; Cheatum, C. M., Line shape analysis of two-dimensional infrared spectra. *J Chem Phys* **2015,** *142* (21), 212427.
17. Johnson, C. A.; Parker, A. W.; Donaldson, P. M.; Garrett-Roe, S., An ultrafast vibrational study of dynamical heterogeneity in the protic ionic liquid ethyl-ammonium nitrate. I. Room temperature dynamics. *The Journal of Chemical Physics* **2021,** *154* (13), 134502.
18. Brinzer, T.; Berquist, E. J.; Ren, Z.; Dutta, S.; Johnson, C. A.; Krisher, C. S.; Lambrecht, D. S.; Garrett-Roe, S., Ultrafast vibrational spectroscopy (2D-IR) of CO2 in ionic liquids: Carbon capture from carbon dioxide's point of view. *The Journal of chemical physics* **2015,** *142* (21), 212425.
19. Ren, Z.; Brinzer, T.; Dutta, S.; Garrett-Roe, S., Thiocyanate as a Local Probe of Ultrafast Structure and Dynamics in Imidazolium-Based Ionic Liquids: Water-Induced Heterogeneity and Cation-Induced Ion Pairing. *The Journal of Physical Chemistry B* **2015,** *119* (13), 4699-4712.

20. Kelsheimer, C.; Garrett-Roe, S., Intramolecular Vibrational Energy Relaxation of CO_2 in Cross-Linked Poly (ethylene glycol) Diacrylate-Based Ion Gels. *The Journal of Physical Chemistry B* **2020,** *125* (5), 1402-1415.

21. Nocedal, J.; Wright, S., *Numerical optimization*. Springer Science & Business Media: 2006.

22. Menke, W., Review of the generalized least squares method. *Surveys in Geophysics* **2015,** *36* (1), 1-25.

23. Marquardt, D. W., An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *J. Soc. Indust. Appl. Math.* **1963,** *11* (2), 431-441.

24. Moré, J. J., The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis*, Springer: 1978; pp 105-116.

25. Fletcher, R., A modified Marquardt subroutine for non-linear least squares. **1971**.

26. Voglis, C.; Lagaris, I. E., Towards "Ideal Multistart". A stochastic approach for locating the minima of a continuous function inside a bounded domain. *Applied Mathematics and Computation* **2009,** *213* (1), 216-229.

27. Robben, K. C.; Cheatum, C. M., Edge-pixel referencing suppresses correlated baseline noise in heterodyned spectroscopies. *The Journal of Chemical Physics* **2020,** *152* (9), 094201.

28. Feng, Y.; Vinogradov, I.; Ge, N. H., General noise suppression scheme with reference detection in heterodyne nonlinear spectroscopy. *Optics Express* **2017,** *25* (21), 26262-26279.

29. Feng, Y.; Vinogradov, I.; Ge, N. H., Optimized noise reduction scheme for heterodyne spectroscopy using array detectors. *Optics Express* **2019,** *27* (15), 20323-20346.

30. Donaldson, P.; Greetham, G.; Shaw, D.; Parker, A.; Towrie, M., A 100 kHz pulse shaping 2D-IR spectrometer based on dual Yb: KGW amplifiers. *The Journal of Physical Chemistry A* **2018,** *122* (3), 780-787.

31. Farrell, K. M.; Ostrander, J. S.; Jones, A. C.; Yakami, B. R.; Dicke, S. S.; Middleton, C. T.; Hamm, P.; Zanni, M. T., Shot-to-shot 2D IR spectroscopy at 100 kHz using a Yb laser and custom-designed electronics. *Optics Express* **2020,** *28* (22), 33584-33602.

32. Kearns, N. M.; Mehlenbacher, R. D.; Jones, A. C.; Zanni, M. T., Broadband 2D electronic spectrometer using white light and pulse shaping: noise and signal evaluation at 1 and 100 kHz. *Optics express* **2017,** *25* (7), 7869-7883.

33. Öztürk, F.; Akdeniz, F., Ill-conditioning and multicollinearity. *Linear Algebra and Its Applications* **2000,** *321* (1-3), 295-305.

34. Stewart, G. W., Collinearity and least squares regression. *Statistical Science* **1987**, 68-84.

35. Dormann, C. F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J. R. G.; Gruber, B.; Lafourcade, B.; Leitao, P. J., Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **2013,** *36* (1), 27-46.

36. Mason, C. H.; Perreault Jr, W. D., Collinearity, power, and interpretation of multiple regression analysis. *Journal of marketing research* **1991,** *28* (3), 268-280.

37. Kim, J. H., Multicollinearity and misleading statistical results. *Korean journal of anesthesiology* **2019,** *72* (6), 558.

38. Seber, G. A. F.; Wild, C. J., *Nonlinear Regression*. John Wiley & Sons, Inc: Hoboken, New Jersey, 2003.

39. Belsley, D. A.; Kuh, E.; Welsch, R. E., *Regression diagnostics: Identifying influential data and sources of collinearity*. John Wiley & Sons: 2005; Vol. 571.

40. Hamm, P.; Zanni, M. T., *Concepts and methods of 2d infrared spectroscopy*. Cambridge University Pres: Cambridge ; New York, 2011; p ix, 286 p.

41. Marple, L., Computing the discrete-time" analytic" signal via FFT. *IEEE Transactions on signal processing* **1999,** *47* (9), 2600-2603.

42. Shim, S.-H.; Strasfeld, D. B.; Zanni, M. T., Generation and characterization of phase and amplitude shaped femtosecond mid-IR pulses. *Optics express* **2006,** *14* (26), 13120-13130.

43. Shim, S.-H.; Zanni, M. T., How to turn your pump–probe instrument into a multidimensional spectrometer: 2D IR and Vis spectroscopies via pulse shaping. *Physical Chemistry Chemical Physics* **2009,** *11* (5), 748-761.

44. Rector, K.; Kwok, A.; Ferrante, C.; Tokmakoff, A.; Rella, C.; Fayer, M., Vibrational anharmonicity and multilevel vibrational dephasing from vibrational echo beats. *The Journal of chemical physics* **1997,** *106* (24), 10027-10036.

45. Yuan, R.; Fayer, M. D., Dynamics of water molecules and ions in concentrated lithium chloride solutions probed with ultrafast 2D IR spectroscopy. *The Journal of Physical Chemistry B* **2019,** *123* (35), 7628-7639.

46. Yuan, R. F.; Yan, C.; Fayer, M., Ion-Molecule Complex Dissociation and Formation Dynamics in LiCl Aqueous Solutions from 2D IR Spectroscopy. *Journal of Physical Chemistry B* **2018,** *122* (46), 10582-10592.

47. Yuan, R. F.; Napoli, J. A.; Yan, C.; Marsalek, O.; Markland, T. E.; Fayer, M. D., Tracking Aqueous Proton Transfer by Two-Dimensional Infrared Spectroscopy and ab Initio Molecular Dynamics Simulations. *Acs Central Science* **2019,** *5* (7), 1269-1277.

48. Falvo, C., A new interpretation of the meaning of the center of line slope from a two-dimensional infrared spectrum. *The Journal of Chemical Physics* **2016,** *144* (23), 234103.

49. Chalmers, J. M., Mid-Infrared Spectroscopy: Anomalies, Artifacts and Common Errors. *Handbook of vibrational spectroscopy* **2006**.

50. Nishida, J.; Tamimi, A.; Fei, H.; Pullen, S.; Ott, S.; Cohen, S. M.; Fayer, M. D., Structural dynamics inside a functionalized metal–organic framework probed by ultrafast 2D IR spectroscopy. *Proceedings of the National Academy of Sciences* **2014,** *111* (52), 18442-18447.

51. Fenn, E. E.; Wong, D. B.; Giammanco, C. H.; Fayer, M., Dynamics of water at the interface in reverse micelles: measurements of spectral diffusion with two-dimensional infrared vibrational echoes. *The Journal of Physical Chemistry B* **2011,** *115* (40), 11658-11670.

52. Hamm, P.; Lim, M.; Hochstrasser, R. M., Non-Markovian dynamics of the vibrations of ions in water from femtosecond infrared three-pulse photon echoes. *Physical review letters* **1998,** *81* (24), 5326.

53. Schmidt-Engler, J. M.; Zangl, R.; Guldan, P.; Morgner, N.; Bredenbeck, J., Exploring the 2D-IR repertoire of the–SCN label to study site-resolved dynamics and solvation in the calcium sensor protein calmodulin. *Physical Chemistry Chemical Physics* **2020,** *22* (10), 5463-5475.

54. Madsen, D.; Stenger, J.; Dreyer, J.; Nibbering, E. T.; Hamm, P.; Elsaesser, T., Coherent vibrational ground-state dynamics of an intramolecular hydrogen bond. *Chemical physics letters* **2001,** *341* (1-2), 56-62.

55. Stenger, J.; Madsen, D.; Dreyer, J.; Nibbering, E. T.; Hamm, P.; Elsaesser, T., Coherent response of hydrogen bonds in liquids probed by ultrafast vibrational spectroscopy. *The Journal of Physical Chemistry A* **2001,** *105* (13), 2929-2932.

56. Nibbering, E. T.; Elsaesser, T., Ultrafast vibrational dynamics of hydrogen bonds in the condensed phase. *Chemical Reviews* **2004,** *104* (4), 1887-1914.

57. Ramasesha, K.; De Marco, L.; Mandal, A.; Tokmakoff, A., Water vibrations have strongly mixed intra- and intermolecular character. *Nature chemistry* **2013,** *5* (11), 935-940.

58. Thämer, M.; De Marco, L.; Ramasesha, K.; Mandal, A.; Tokmakoff, A., Ultrafast 2D IR spectroscopy of the excess proton in liquid water. *Science* **2015,** *350* (6256), 78-82.

59. Lock, A.; Woutersen, S.; Bakker, H., Ultrafast energy equilibration in hydrogen-bonded liquids. *The Journal of Physical Chemistry A* **2001,** *105* (8), 1238-1243.