Supporting Information For

CLS Next Gen: Accurate Frequency-Frequency Correlation Functions from Center Line Slope Analysis of 2D Correlation Spectra using Artificial Neural Networks

David J. Hoffman and Michael D. Fayer*
Department of Chemistry
Stanford University, Stanford, CA 94305
*Phone: 650 723-4446; Email: fayer@stanford.edu

Table of Contents

9 Pages

1 Figure

1 Table

1. Complete Neural Network Calculations

For an FFCF parameter neural network discussed in the main text with inputs x and output y, the explicit calculations to go from x to y are as follows:

$$X(i) = \left(x(i) - x_{off}(i)\right)x_{gain}(i) - 1 \tag{S1}$$

$$p_1(j) = \tanh\left(b_1(j) + \sum_{i=1}^{5} M_1(i,j)X(i)\right) \tag{S2}$$

$$p_2 = b_2 + \sum_{i=1}^{n} M_2(i)p_1(i) \tag{S3}$$

$$y = (p_2 + 1)/y_{gain} + y_{off} \tag{S4}$$

For all three networks, the inputs x are:

$$\begin{aligned}
x(1) &= A_1 \\
x(2) &= \log_{10}(s_1) \\
x(3) &= A_2 \\
x(4) &= \log_{10}(s_2) \\
x(5) &= A_\infty
\end{aligned} \tag{S5}$$

where A_∞ is the amplitude of a time-independent component, or offset, of the CLS if there is one. The values of $\log(s)$ become singular when $s = 0$, so in the case of $A_1 = 0$, $x(2)$ were arbitrarily defined to be +4 and when $A_2 = 0$, $x(4)$ was defined to be +6.7. The inputs are modified in the case of a triexponential or triexponential to an offset CLS decay, as discussed in the main text and in the following section.

The ANN parameters x_{off}, x_{gain}, b_1, M_1, b_2, M_2, y_{off}, and y_{gain} are listed for the three networks in the Supporting Information Excel spreadsheet (each ANN has its own sheet).

2. Complete Calculations for FFCF Delta Parameters

As was discussed in the main text, we parameterize the CLS as follows:

$$\mathrm{CLS}(T_w) = \sum_i A_i \exp(-T_w / \tau_i), \tag{S6}$$

from which we can derive the rescaled CLS time scales as:

$$s_i = \mathrm{FWHM} \cdot \left(A_i / \sum_j A_j \right)^{1/2} \tau_i, \tag{S7}$$

where FWHM is the full width of the linear absorption spectrum in units of angular frequency

(e.g., radians/ps) and the τ_i are in ps.

The CLS amplitudes and rescaled time components are input to the neural networks to

get the associated FFCF σ terms. For CLS terms with up to two decay terms and with or without

an offset (i.e., single exponential, single exponential to an offset, biexponential, or biexponential

to an offset), the calculations are as follows:

$$\begin{aligned}
\log_{10}(\sigma_1) &= f_1\left(A_1, \log_{10}(s_1), A_2, \log_{10}(s_2), A_\infty \right) \\
\log_{10}(\sigma_2) &= f_2\left(A_1, \log_{10}(s_1), A_2, \log_{10}(s_2), A_\infty \right), \\
\sigma_\infty &= f_\infty\left(A_1, \log_{10}(s_1), A_2, \log_{10}(s_2), A_\infty \right)
\end{aligned} \tag{S8}$$

where f_i correspond to the neural networks for σ_i defined in the previous section and in the

Supporting Information Excel spreadsheet. For a triexponential or triexponential to an offset, the

modification can be applied as was described in the main text:

$$\begin{aligned}
\log_{10}(\sigma_1) &= f_1\left(A_1, \log_{10}(s_1), A_2, \log_{10}(s_2), A_3 + A_\infty \right) \\
\log_{10}(\sigma_2) &= f_2\left(A_1, \log_{10}(s_1), A_2, \log_{10}(s_2), A_3 + A_\infty \right) \\
\log_{10}(\sigma_3) &= f_2\left(A_1 + A_2, \log_{10}\left(\sqrt{s_1 s_2}\right), A_3, \log_{10}(s_3), A_\infty \right) \cdot \\
\sigma_\infty &= f_\infty\left(A_1, \log_{10}(s_1), A_2 + A_3, \log_{10}\left(\sqrt{s_2 s_3}\right), A_\infty \right)
\end{aligned} \tag{S9}$$

As the τ_i terms are conserved between the FFCF and the CLS, the Δ_i can be obtained for

the standard σ_i from a simple calculation:

$$\Delta_i = \sigma_i / \tau_i, \tag{S10}$$

which returns the Δ_i as an angular frequency (e.g. radians/ps). The offset term Δ_∞ can similarly be constructed from the σ_∞ and the linear absorption spectrum FWHM:

$$\Delta_\infty = \frac{\text{FWHM}}{2\sqrt{2\ln 2}}\sqrt{\sigma_\infty},$$ (S11)

where Δ_∞ will have the same units as the input FWHM term.

3. Complete Calculations for FFCF Homogeneous Line Width

Once the Δ's are calculated, the only remaining unknown parameter in the FFCF is the homogeneous line width, $\Gamma = (\pi T_2)^{-1}$. To determine Γ, we will need to calculate the FWHM of the linear absorption line shape from the FFCF for an arbitrary Γ.

From the definition of the FFCF (Eq. 5) input into the definition of the line shape function, $g(t)$, (Eq. 3) we get for a generic FFCF:

$$g(t) = t/T_2 + \sum_i \Delta_i^2 \tau_i t + \Delta_i^2 \tau_i^2 (\exp(-t/\tau_i) - 1),$$ (S12)

where the Δ_i have units of angular frequency. Eq. S12 can then be input into the linear response function:

$$R^1(t) = \exp(-g(t)).$$ (S13)

Eq. S13 is then Fourier transformed to obtain the absorption spectrum line shape.

For numerical calculations, first a t_f was found that gave $R^1(t_f) < .0005$ to minimize apodization of the line shape. The interval from 0 to t_f was then discretized into 100 equally spaced time steps at which R^1 was evaluated. The discretized response function was zero-padded to a total of 512 time steps to increase the frequency resolution and fast Fourier transformed to yield the linear spectrum. The FWHM of the spectrum was calculated by interpolating the resulting line shape and finding the frequencies where the amplitude of the spectrum was half of

the maximum. This procedure for calculating the FWHM from a given FFCF will be denoted as $\Lambda(T_2; \Delta_1, \Delta_2, \ldots; \tau_1, \tau_2, \ldots)$.

Before attempting to calculate the experimental FFCF, calculating Λ for a very large or infinite T_2 can be done to make sure that Λ is less than the experimental FWHM. This validates the calculations that determine the Δ parameters and checks that there is a solution to the minimization problem.

Once it is known that a solution exists, the goal is then to minimize the function:

$$h(T_2) = \left(\Lambda\left(T_2; \Delta_1, \Delta_2, \ldots; \tau_1, \tau_2, \ldots\right) - \mathrm{FWHM} \right)^2 \tag{S14}$$

which will yield the T_2 that gives the correct FWHM. The algorithm described above for Λ does not have easily calculable derivatives that could be used for gradient descent algorithms, so a non-gradient simplex solver was used instead (fminsearch in the MATLAB software was used). This procedure is then capable of giving results within the precision of the calculation of Λ.

4. Neural Network Standard Error and Error Propagation from CLS Fits

As can be seen in Fig. 7 of the main text, the ANNs provide high quality function approximations, but they are not exact. For the regressions in Fig. 7, the residuals have the statistics shown in Table S1. It should be noted that some of the error arises from computational errors in the training data themselves, so the reported standard deviations include noise from both the computational data and the neural network approximations themselves, and does not account for systematic errors that arose during the calculations of the training data. The residual is also typically larger with increased motional narrowing (See Fig. S1); only the total variance reported in Table S1 is used for these calculations.
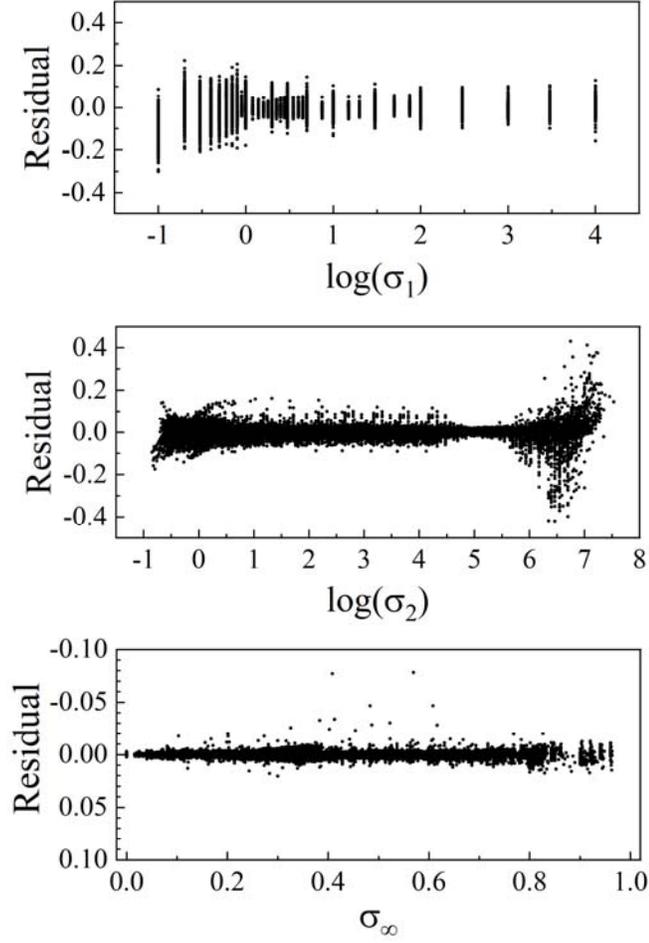
Fig. S1: Residuals for neural networks relative to their training data. Statistics in Table S1.

Table S1: Neural Network Residuals Statistics. Errors relative to the calculated training data

Network	Standard Error	Maximum Residual
$\log(\sigma_1)$	0.01452	0.29999
$\log(\sigma_2)$	0.01173	0.57731
σ_∞	0.00138	0.07817

The variance of the neural network approximation can be propagated through to the

corresponding Δ terms using standard methods:

$$\mathrm{Var}\left(\Delta_i\right) = \left(\frac{\delta\Delta_i}{\delta\sigma_i}\right)^2 \mathrm{Var}\left(\sigma_i\right). \tag{S15}$$

Assuming that the errors between different networks are uncorrelated, the variance in the calculation of the homogeneous line width is:

$$\mathrm{Var}\left(\Gamma\right) = \sum_i \left(\frac{\delta\Gamma}{\delta\Delta_i}\right)^2 \mathrm{Var}\left(\Delta_i\right). \tag{S16}$$

The gradients of Γ are calculated numerically in the neighborhoods of the algorithmically determined Δ_i parameters. The systematic errors of the calculation of Γ was not considered.

For experimental data, there will also be error in the CLS fit parameters and the FWHM of the absorption line shape. The CLS fit parameters are also highly covariant with each other: for instance, if one amplitude increases, other amplitudes will tend to decrease. To properly accommodate these correlations, the complete covariance matrix is required. Including the intrinsic error of the neural network calculations and the extrinsic error from the experimental parameters, we get for a generic FFCF parameter q and experimental parameters B:

$$\mathrm{Var}\left(q(B)\right) = \mathrm{Var}\left(q(\beta)\right) + \nabla q(\beta)^T \cdot \mathrm{Cov}(B) \cdot \nabla q(\beta). \tag{S17}$$

Here β is the "true value" of B, which is the value which B is normally distributed around under the central limit theorem. In practice, β is the actual fit value of the experimental parameters, neglecting error. $\mathrm{Cov}(B)$ is the covariance matrix of the experimental parameters, and $\nabla q(\beta)$ is the gradient of relevant FFCF parameters with respect to the experimental parameters. As in Eq. S16, these gradients are calculated numerically in the neighborhood of β. The variance of an FFCF parameter is then the sum of the intrinsic variance introduced by the neural networks $\mathrm{Var}(q(\beta))$ and the variance introduced by uncertainty in the experimental parameters.

Generally the error from the fit parameters was seen to be much larger than the intrinsic error from the neural networks. Additionally, these calculations demonstrate that the homogeneous line width Γ is the term with the highest uncertainty in these calculations. The Δ

terms principally have errors introduced by the intrinsic neural network variance, the error in the corresponding CLS amplitude, and, if the component is motionally narrowed, in the variance of the time constant. However, the Γ terms depend on every Δ term, which means it has error that arises from each neural network calculation (Eq. S16), each CLS amplitude, and each time constant for a motionally narrowed component. The end result is much higher relative error in the Γ term than the Δ terms.

5. Calculations of a Quantum FFCF in Various Limits

In Section V. C of the main text, the results of modifying a classical, motionally narrowed FFCF with two different quantum FFCF phenomena were presented. In particular, two limiting cases that are of importance for 2D ES as indicated by Šanda et Al:[1] the Stokes shift from the overdamped limit and the high frequency coupling from the underdamped limit. For the Stokes shift, the high temperature overdamped limit was implimented with the following complex-valued line shape function for an FFCF component:

$$g(t) = \tau^2 \left(\frac{2kT\lambda}{\hbar} - i\frac{\lambda}{\tau} \right)\left(e^{-t/\tau} + t/\tau - 1 \right). \tag{S18}$$

Where λ is the coupling strength/magnitude of the Stokes shift. The line shape function can then be used in the response functions to calculate the absorption spectrum or 2D spectra as usual. Using:

$$\Delta^2 = \frac{2\lambda kT}{\hbar}, \tag{S19}$$

it is clear that in the low coupling/high temperature limit, Eq. S18 goes to the classical Gaussian process line shape function (Eq. 6). However, Eq. S18 is itself a (weaker) high temperature approximation, and is only valid in the regime that $kT > \hbar/\tau$.

For the model calculations in Section V. C, the highly motionally narrowed FFCF of the CN stretch of MeSCN in H_2O was used, which has parameters ($\Delta_1 = 7.4$ cm^{-1}, $\tau_1 = 0.4$ ps, $\Delta_2 = 2.6$ cm^{-1}, $\tau_2 = 1.7$ ps, $T_2 = 1$ ps). The temperature condition depends on the fastest time constant, so the lowest temperature that can be examined in this limit is $T = \hbar/k\tau_1 = 20$ K. From this calculation, the greatest coupling strength, λ, could be determined for each line shape component by putting in T and Δ_1 or Δ_2 into Eq. S19 and solving for λ. This procedure gives $\lambda_1 \approx 0.3\ \Delta_1$ and $\lambda_2 \approx 0.1\ \Delta_2$. The results for $T = 20$ K were presented in the main text.

For the underdamped limit, it was assumed that the oscillatory part and the diffusive part of the response function were separable as:

$$R_{tot}(t_1, T_w, t_3) = R_{dif}(t_1, T_w, t_3)R_{vib}(t_1, T_w, t_3) \tag{S20}$$

For a total response function R_{tot} that is made up of a classical, diffusive component R_{dif} (again using the CN stretch of MeSCN in H_2O parameters for the model calculation) and an oscillatory part R_{vib}. The line shape function for the underdamped, vibrational part is then:

$$g(t) = \frac{2\lambda}{\Omega}\coth\left(\frac{\hbar\Omega}{2kT}\right)\left(1 - e^{-\gamma t}\cos\Omega t\right) + i\frac{2\lambda}{\Omega}e^{-\gamma t}\sin\Omega t - 2i\lambda t \tag{S21}$$

Where λ is still a coupling parameter, γ is the damping rate, and Ω is the oscillation frequency, which is much greater than γ. For the model calculations shown in the main text, $T = 300$ K, $\lambda = 0.75$ ps^{-1}, $\Omega = 78$ ps^{-1}, and $\gamma = 0.5$ ps^{-1}. The effect of motional narrowing in the diffusive part was found to be totally independent of the choice of parameters.

SI References

1.   Šanda, F.; Perlík, V.; Lincoln, C. N.; Hauer, J., Center Line Slope Analysis in Two-Dimensional Electronic Spectroscopy. *J. Phys. Chem. A* **2015,** *119*, 10893-10909.