A New Method Based on Pseudo-Zernike Polynomials to Analyze and Extract Dynamical and Spectral Information from the 2DIR Spectra

*Anit Gurung and Daniel G. Kuroda**

Department of Chemistry, Louisiana State University, Baton Rouge, Louisiana 70803, USA

*Address corresponds to dkuroda@lsu.edu

ABSTRACT

Ultrafast two-dimensional infrared (2DIR) spectroscopy is a relatively new methodology, which has now been widely used to study the molecular structure and dynamics of molecular processes occurring in solution. Typically, in 2DIR spectroscopy the dynamics of the system is inferred from the evolution of 2DIR spectral features over waiting times. One of the most important metrics derived from the 2DIR is the frequency-frequency correlation function (FFCF), which can be extracted using different methods, including center and nodal line slope. However, these methods struggle to correctly describe the dynamics in 2DIR spectra with multiple and overlapping transitions. Here, a new approach, utilizing pseudo-Zernike moments, is introduced to retrieve the FFCF dynamics of each spectral component from complex 2DIR spectra. The results show that this new method not only produces equivalent results to more established methodologies in simple spectra, but also successfully extracts the FFCF dynamics of individual component from very congested and unresolved 2DIR spectra. In addition, this new methodology can be used to locate the individual frequency components from those complex spectra. Overall, a new methodology for analyzing the 2D spectra is presented here, which allows us to retrieve previously unattainable spectral features from the 2DIR spectra.

INTRODUCTION

Two-dimensional infrared (2DIR) spectroscopy is a powerful and well-established time resolved laser spectroscopy. 2DIR spectroscopy retrieves the dynamics of the thermal motions occurring in liquid systems at picosecond timescales. To this end, the spectroscopy, by means of femtosecond infrared pulses, creates and follows the vibrational coherences in the sample in order to obtain the time scale of the decorrelation of the coherences.[1] Different molecular processes can cause the observed decorrelations, with spectral diffusion being one of the most important. Spectral diffusion is defined as the change in vibrational energy level of an oscillator due to fluctuations in its molecular environment.[1] Hence, the spectral diffusion process provides insights into the time-dependent molecular interactions exerted by the surroundings on the molecular system.[1] This process is usually modeled in terms of the so-called frequency-frequency correlation function (FFCF), which contains the amplitude and characteristic time of the frequency fluctuations.[1] Experimentally, the FFCF dynamics is obtained either by fitting the 2DIR spectra[2-4] or by measuring different 2DIR spectral metrics as function of waiting time.

Several FFCF extraction methods based on different 2DIR spectral features and time components (i.e., center line slope (CLS),[5, 6] nodal line slope (NLS),[7, 8] slope of the phase,[9] inhomogeneity index,[9, 10] eccentricity,[11] peak shift,[12-14] ellipticity,[15, 16] and dynamic line width[17, 18]) have been previously used for this purpose. The vast majority of these methods utilize specific spectral features of the 2DIR peak to retrieve the dynamics of the FFCF. However, each technique has its own practical applications, advantages and limitations in the analysis of 2DIR spectra. For instance, most methods efficiently extract the FFCF characteristic times from spectra containing a single transition within the investigated IR region. In this case, the CLS and NLS analysis have been proven to be the most consistent, reliable and simple among all methods for obtaining the

FFCF dynamics for 2DIR spectra with a single transition.[5-8] However, most methods either fail or require a more complex analysis when dealing with complex 2DIR spectra containing multiple transitions within the studied window.[19]

The complexity of 2DIR spectra containing multiple transitions increases significantly due to differences in vibrational anharmonicities of the transitions, overlap of the vibrational bands and the presence of cross peaks. In this complex scenario, even established methods like CLS and NLS are not particularly useful to retrieve the FFCF dynamics of system.[7, 19] Note that the CLS method can be used to extract the FFCF dynamics from the 2DIR spectra containing two transitions, but prior knowledge of the FFCF dynamics for one of the components is required.[19] Finally, the results of the CLS method strongly depend on the modeling used for retrieving the metric, as well as the quality of the data.[5]

In this article, a new method based on pseudo-Zernike moments (PZM)[20, 21] is introduced to analyze the underlying dynamics from the 2DIR spectra. PZM is based on the description of 2D data using pseudo-Zernike polynomials (PZP).[22] The main properties of PZM are rotation nvariance, and translation/scale variance.[20, 21] In particular, PZMs have shown to suffer from less redundancy and noise sensitivity when analyzing 2D functions due to their orthogonality property of the PZPs in the unit circle.[23] Hence, PZMs have been successfully used for feature extraction methodology in pattern recognition.[24, 25] Examples of the applications of PZMs are the characterization of the wave front aberration of both human eyes and optical systems.[26, 27] In addition, the time evolution of PZMs has been previously used for studying the time progression of eye diseases, such as age-related macular degeneration.[28] Hence, the similarity of the circular/oval nature of PZP components when compared to the 2DIR spectral features provides a suitable framework for the description of the 2DIR spectra and their waiting time evolution. The

adequacy of the method to reproduce the 2DIR spectra is demonstrated by applying the PZP decomposition to a 2DIR spectrum (Figure 1). Graphically, the recovered 2DIR spectra as a function of the PZP order show that the first 15th orders are sufficient to reproduce a 2DIR spectrum in agreement with the error of the modeling from the RMSD (see Supplemental Material).

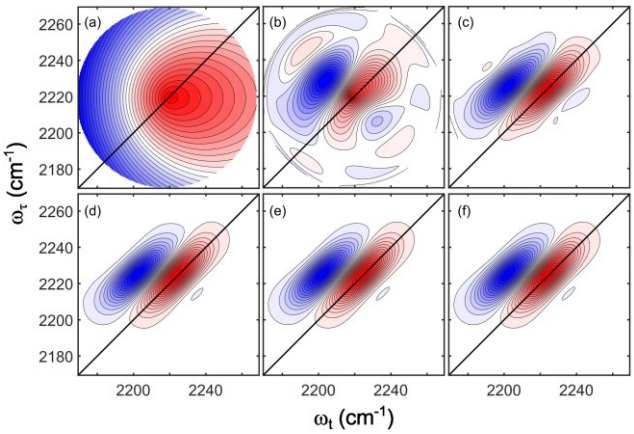

Figure 1. Recovered 2DIR spectra using PZPs up to 1st, 5th, 10th, 15th, and 20th order PZM represented with panel (a), (b), (c), (d) and (e) respectively. Panel (f) corresponds to the original 2DIR spectrum.

The PZM decomposition provides a unique representation of the 2DIR spectral features, such as 2D peak shape, in terms of orthogonal pseudo-Zernike polynomials.[20] This PZP, or equivalent PZM, representation can be applied to the retrieval of temporal and spectral information from of the waiting time evolution of the 2DIR spectra even in complex situations, such as those created by the presence of overlapping and unresolved transitions. In addition, the PZM methodology does not require any prior knowledge of the contributions from individual components in the 2DIR spectra. Overall, the new PZM method is a complementary approach to CLS and NLS for in-depth analysis of 2DIR spectra, especially when the 2DIR spectra contain information from multiple vibrational components.

THEORY

Pseudo-Zernike Moments

The Pseudo-Zernike polynomials (PZP) are based on the polynomials first introduced by Frits Zernike in 1934.[29] PZPs comprise a set of linearly independent polynomials arising from complex-valued exponential and real-valued radial polynomial functions in polar coordinates.[21] Mathematically, the PZPs are expressed as:

$$V_{nm}(x,y)=V_{nm}(r,\theta)=R_{nm}(r)e^{im\theta} \qquad (1)$$

where n and m are integers that fulfil $|m| \leq n$, and r and θ are normalized polar coordinate variables (i.e., $|r| \leq 1$ and $0 \leq \theta \leq 2\pi$). Due to the completeness and orthogonality nature of the PZPs, these polynomials can be used to represent 2D functions (i.e., f(x, y)) by projecting the function into the PZP basis set, such that the function is represented as

$$f(x,y)=\sum_n \sum_m A_{nm} V_{nm}(r,\theta) \qquad (2)$$

where the A_{nm} are the coefficients. The coefficients (A_{nm}) of the PZP projection, so-called Pseudo-Zernike moments (PZM), are calculated as follows:

$$A_{nm}=\frac{n+1}{\pi} \sum_x \sum_y f(x,y)\ V^*_{nm}(r,\theta),\ x^2+y^2 \leq 1 \qquad (3)$$

It is important to note that the orthogonality of the PZPs results in almost zero redundancy in their moments resulting in a direct correspondence between the different moments and the distinct features of the 2D function.[21]

The PZP basis set is large since for up to an n^{th} degree, the PZP basis have $(n+1)^2$ polynomials. The large size of this basis sets is advantageous for image representation due to their high feature

extraction capability and robustness in the presence of noise.[20] However, the large number of PZPs also restricts the computation of PZMs to small sections of the space due to its high computational cost.[30]

Similarity Measure

The large dimensionality of the PZP space provides an intractable number of PZMs (coefficients) when describing any arbitrary 2D function. In particular, for the case of the 2DIR spectra, each spectrum, corresponding to a given waiting time, can require a large set of PZMs for their accurate representation. Moreover, it is expected that the time evolution of the 2DIR spectra directly relates to changes of the PZMs. However, individual PZMs cannot be easily associated to the time evolution of the specific spectral features of the 2DIR spectra due to the large number of coefficients typically needed for their representation. For example, in the previously reproduced spectra (Figure 1) 256 PZPs are need. To overcome this issue, the cosine similarity (CS) measure among PZPs is used here. The CS represents the angle between two vectors in the inner product space and is one of the most common similarity measures for high dimensional vector spaces.[31-33] The CS between two vectors, $\mathbf{p}_1$ and $\mathbf{p}_2$, is mathematically defined as follows:

$$\cos\theta = \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{\|\mathbf{p}_1\|\|\mathbf{p}_2\|}$$

where θ is the angle between the two vectors in the inner product space. Physically, the CSs represent how similar two data sets (spectra) are. For example, two vectorized data sets containing the same information will have a CS close to one showcasing their strong similarity. In contrast, data sets with CS values close to zero are considered to be very dissimilar (i.e., contain very

different information). Hence, the calculation of the CS can be used to quantify the similarity of the PZMs when describing two different 2DIR spectra.

Inherently, the CS provides a relative comparison of the analyzed vectors Hence, the PZMs derived for the 2DIR spectrum of a given waiting time have two possible similarity factors arising from its comparison with the PZMs of either the initial waiting time (t_0), which is typically Tw=0 ps, or the final waiting time (t_f) corresponding to 2DIR spectrum of the longest waiting time used for the time dependence analysis. The first similarity factor, CS_{tw}, relates to the angle changes of the cosine projection with respect to the PZMs of the initial 2DIR spectrum, and the second, cosine distance (CD_{tw}), to the 2DIR spectrum at final waiting time. The CS_{tw} and CD_{tw} are expressed as follows:

$$CS_{t_w} = \cos\theta = \frac{\boldsymbol{PZM}_{t_0} \cdot \boldsymbol{PZM}_{t_w}}{\left\|\boldsymbol{PZM}_{t_0}\right\| \left\|\boldsymbol{PZM}_{t_w}\right\|}$$

$$CD_{t_w} = 1 - \cos\theta = 1 - \frac{\boldsymbol{PZM}_{t_f} \cdot \boldsymbol{PZM}_{t_w}}{\left\|\boldsymbol{PZM}_{t_f}\right\| \left\|\boldsymbol{PZM}_{t_w}\right\|}$$

where t_w represents the waiting of the similarity factor between either the initial waiting time (t_0 for CSt_w) or the final time (t_f for CD_{tw}) considered in the analysis. Note that these two similarity factors are defined such that their limits are 1 and 0 for the initial and final times of CStw and CDtw, respectively, since they represent how similar the analyzed 2DIR spectrum is with respect to its corresponding waiting time spectrum. It is important to note that the 2DIR spectra are very similar to each other regardless of their origin. It is therefore unlikely to see that the similarity factors between a pair of 2DIR spectra will be zero. In summary, the CSs only quantify how similar the two spectra are, not how much they can possibly be as the CS is not absolute, but relative.

Therefore, dynamics on much longer time scales than the investigated window will not be captured from the 2DIR spectra by either CS due to the relative nature of the analysis.

The time evolution of the 2DIR spectra as a function of waiting time produces a gradual change in both similarity factors (CS_{tw} and CD_{tw}). The high (low) values for CS_{tw} (CD_{tw}) showcases the correlation of 2D spectral features seen in the 2D spectrum at given waiting time with respect to the initial (final) spectral features. Therefore, one can use these two CS measures to extract the FFCF and its dynamical time scale from the 2DIR spectra. However, both similarity factors are biased due to significant dependence on the individual PZMs to the reference PZM vector. To minimize bias, a linear combination of the similarity factors (CS_{tw} and CD_{tw}) is introduced, which is referred here as pseudo-Zernike similarity (PZS). The computation of PZSs is carried out within the cosine subspace, where CS_{tw} and CD_{tw} are assumed to be orthogonal. This assumption is consistent with the observation that neither the initial nor the final 2DIR spectra in the time series are sufficient to represent the waiting time evolution of the 2DIR spectra. The mathematical form of PZS is:

$$PZS_{t_w} = (cos\,45°)CS_{t_w} + (1 - cos\,45°)CD_{t_w}$$

where the two scalar parameters (i.e., $(cos\,45°)$ and $(1 - cos\,45°)$) arise from assuming that there is equal contribution of the similarity factors representing the initial and final 2DIR spectra, or mathematically, midpoint between orthogonal vectors. Unlike the CS_{tw} or CD_{tw}, the PZS as function of waiting time (PZS_{tw}) provides a metric that directly estimates the dynamics of the FFCF from the 2DIR spectrum (see Supplementary Material).

METHODS

Software

The analysis of the data was performed using MATLAB 2022.[34] The PZS analysis code is based on the publicly available code of Pseudo-Zernike Functions.[35] The PZS analysis code can be found free of charge at https://github.com/dkurodalab/PZSanalysis.

PZS analysis

The PZS analysis is performed in three different ways: the total spectrum analysis (TPZS), the frequency resolved analysis (FR-PZS), and the PZS analysis along the center line (CL-PZS).

I. Total PZS (TPZS)

The TPZS analysis is performed by selecting a particular region of the 2DIR spectra and computing the PZS at each waiting time (i.e., PZS_{tw}). Hence, the TPZS is aimed at analyzing the most important region of the 2DIR around the maximum with low spectral resolution. In particular, this analysis involves using the circular area positioned around the overall maximum of the 2DIR spectrum at initial and final waiting time. To this end, a circle, along both ω_τ and ω_t directions, with a radius containing the full width half maximum of the diagonal trace at the initial waiting time is selected for doing the analysis. It is important to note that the peaks in the 2DIR spectra can undergo spectral shift as function of waiting time due to the presence of other signals (such non-resonant from the solvent or cell materials) or other experimental uncertainties. Therefore, the effect is removed by centering the analysis area with respect to the maximum of the peak at any given waiting time. There are many n^{th}-order PZMs that can generate a feature vector accurately describing the unique features of the selected region. Hence, the feature vector comprised of 0 through n^{th} order PZMs is derived from a convergence plot computed during PZS calculations (see

Supplementary Material). For consistency, the size of feature vector is kept fixed at 15th order PZM on all studied presented in this work.

II. Frequency resolved PZS (FR-PZS)

Another approach to analyze the 2DIR spectra is by frequency resolving the PZS analysis (FR-PZS). Unlike TPZS, the FR-PZS utilizes a rectangular spectral region with small ω_τ and large ω_t frequency window enclosed within a circle that contain the selected spectral region. Hence, this approach is suitable for frequency resolving the PZS analysis within a section of the 2DIR spectra at the expense of experiencing greater uncertainty in the dynamics. Application of this methodology leads to a scan of the PZS along the excitation frequency (ω_τ). Due to the number of PZS time evolutions produced in this analysis, the changes in the PZS dynamics for each ω_τ window are modeled with a function of the form: $y_0 + Ae^{-\frac{t_w}{\tau}}$, where y_0 is the PZM similarity factor at high waiting times, and A is the amplitude of the decay time and τ is the characteristic time of the decay function. However, in systems with known biexponential decays, an average of the correlation time is used.[36]

III. PZS through center line (CL-PZS)

CL-PZS analysis is performed by centering the TPZS analysis at frequencies along the center line, where the CLS at a given excitation frequency (ω_τ) acts as a centroid for analysis area in 2DIR spectrum. The TPZS and CL-PZS analysis differ because the latter analyzes a significantly larger frequency window along ω_τ than the former, but the TPZS is the particular case of the CL-PZS at the maximum. The large window investigated in the CL-PZS analysis gives weight to different spectral regions than those considered in the TPZS analysis, which is particularly important when using the method on 2DIR spectra with more than one transition. This is particularly important

because, even in complex 2DIR spectra, the different spectral regions have the same waiting time evolution, which provides redundancy to validate PZS methodology.

DATA SETS

Two different data sets were selected for the analysis: simulated and experimental 2DIR spectra. Representative 2DIR spectra for each data set can be found in the Supplementary section.

I. Simulated 2DIR spectra

The 2DIR spectra was simulated using the response function approach previously described in the literature.[37] For a single component system, the parameters are similar to those derived from benzonitrile[38] and include: central peak frequency (ω_0), anharmonicity ($\Delta\omega$), vibrational lifetime (T_{10}), transition dipole moment (μ), and a FFCF containing a single Kubo function with an amplitude of the frequency fluctuation (Δ) and time decay (τ). The specific parameters of different single component systems studied here are shown in Table 1. In addition, cases where the 2DIR spectra have more than one component were also simulated. Since most 2DIR spectra containing more than one transition can be reduced to a system of two components, the simulation of multicomponent contained only two components with different frequencies (ω_1 and ω_2) and different FFCF dynamics (τ_1 and τ_2). Two cases were studied where the selected center frequencies lead to either resolved or unresolved peaks in the 2DIR spectra (Table 1).

Table 1. Parameters used in the simulated 2DIR spectra.

Number of components	Case	ω_0 (cm^{-1})	$\Delta\omega$ (cm^{-1})	T_{10} (ps)	μ	Δ (cm^{-1})	τ_{FFCF} (ps)
1	I	2225	22	4.2	1	2	3
1	II	2225	22	4.2	1	2	6
2	III	2215 2235	22	4.2	1	2	6 3
2	IV	2220 2230	22	4.2	1	2	6 3

II. Experimental 2DIR Spectra

Five different experimental datasets are used in this study. The first three cases represent simple 2DIR spectra containing a single vibrational transition in the investigated window, while the other two cases correspond to more complex 2DIR spectra arising from multiple vibrational transitions. For the single transition, the selected 2DIR spectra datasets correspond to those probing the asymmetric stretch of aqueous azide ion,[39, 40] as well as the nitrile stretch of benzyl thiocyanate (BzSCN) in a deep eutectic solvent (DES).[41, 42] While the two systems contain a single vibrational transition in the different spectral regions, the nitrile stretch in the organic thiocyanate has considerably larger anharmonicity,[43] which allow us to study the effect of anharmonicity on the PZS analysis. Finally, the third system also consisted of investigating the N3 asymmetric of the azide ion, but in a very diluted solution allowing us to investigate the effect of noise in PZS analysis.

The other two set of dataset correspond to those probing the amide I mode of an aqueous solution of N-isopropyl propionamide and the nitrile stretch of the thiocyanate ion in an amide-based DES.[43, 44] While the first case correspond to a system with two well-resolved vibrational transitions in the 2DIR spectra, the second case provide an example of unresolved in which the presence of two transition is only derived from the shape of the 2DIR spectra due to the presence of crosspeaks. Therefore, these two sets of complex 2DIR spectra are useful cases for studying the viability of PZS analysis in the presence of intricate and time dependent 2D spectral features.

RESULTS AND DISCUSSION

Single Transition

Due to the possible complexity of the PZS analysis, the suitability of the PZS method for extracting the FFCF dynamics out of the waiting time evolution of the 2DIR spectra was first investigated using single component systems in both experimental and simulated 2D spectra. In the simulated case, the 2DIR spectra consisted of a single transition with a FFCF having a single exponential decay function with two possible decay times (case I and II of Table 1). The TPZS analysis on either case produces a CS factor with a time dependence (Figure 2), which is well described by an exponential decay of the form: $y_0 + Ae^{-\frac{t_w}{\tau_{FFCF}}}$, where y_0 is the minimum PZS factor, A is the amplitude of the decay time and τ_{FFCF} is the characteristic time of the FFCF decay function. The FFCF time constants retrieved from the modeling of the time dependent PZS for case I and II are in very good agreement with the values used for the simulation (Table 2). This last result is very important because it demonstrates the applicability of the TPZS analysis to retrieve the FFCF dynamics from the time evolution of the 2DIR analysis.

A comparison between TPZS analysis and any of the two most traditional methods (i.e., NLS and CLS) shows that the new methodology performs comparable to the other two more established methods. Specifically, the error associated with τ_{FFCF} from TPZS analysis is similar or lower than CLS, but slightly higher than NLS. A better view of the TPZS performance versus other more established methods is obtained from the correlation plot between the similarity factor (PZS_{tw}) and the metrics derived from the traditional methods (insets of Figure 2), because it does not require modeling the data. As expected, both plots show a strong linear correlation ($R^2 > 0.99$ for any of the cases), which further support the usability of TPZS analysis for retrieving the FFCF dynamics from the 2DIR spectra. It also observed from the TPZS analysis that there is non-zero offset, which does not match that of the CLS or NLS analysis. This offset in the TPZS analysis arises from the

use of CS factors (see methods section) and hinders the possibility of determining the presence of a static inhomogeneous contribution in the 2DIR spectra with the TPZS method alone.

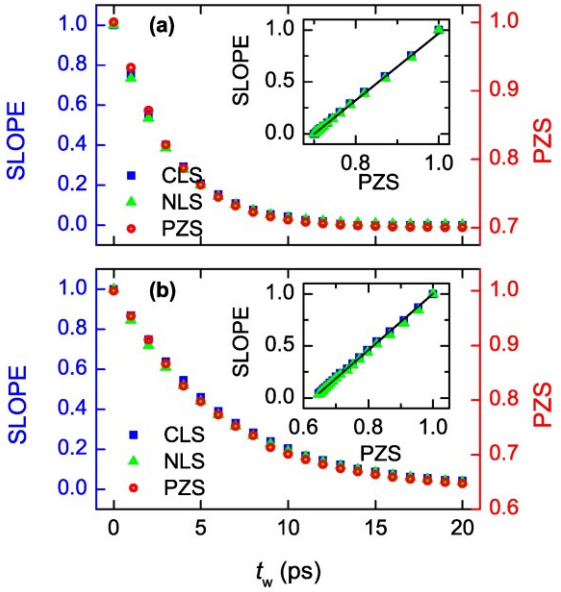

Figure 2. Comparison among FFCF retrieval methods for simulated data. Panels (a) and (b) show analysis of case I (t=3ps) and case II (t=6ps) of the simulations using CLS (solid blue square), NLS (solid green triangle) and TPZS (hollow red circle).

Table 2. PZS, CLS and NLS analysis of simulated data. Characteristic time of the dynamics derived from the analysis of simulated 2DIR spectra containing one or two vibrational components.

Components	τ (ps)					
	Theory	TPZS	FR-PZS	CL-PZS	CLS	NLS
One	3.00	3.28 ± 0.06	3.48 ± 0.08	--	3.27 ± 0.03	3.11 ± 0.01
	6.00	6.2 ± 0.1	6.8 ± 0.2	--	6.38 ± 0.06	6.08 ± 0.01
Two (resolved)	τ_1 = 6.00	--	5.89 ± 0.02	6.6 ± 0.3	10.4 ± 1.2	7.6 ± 0.4
	τ_2 = 3.00	--	3.28 ± 0.05	3.6 ± 0.1	4.5 ± 0.3	4.7 ± 0.2
Two (unresolved)	τ_1 = 6.00	--	6.8 ± 0.2	6.4 ± 0.3	4.49 ± 0.06*	4.09 ± 0.02*
	τ_2 = 3.00	--	3.60 ± 0.02	3.3 ± 0.1		

* These results are average values because neither analysis method can separate the two components.

The new PZS method was also tested using experimental 2DIR spectra of the three samples having a single transition. In the 50mM azide aqueous solution, the FFCF evolution obtained using the TPZS analysis produces a decaying CS factor (Figure 3), which is well modeled with single exponential decay function. The characteristic time derived from CS time evolution is comparable

to the NLS and CLS values, and previous experimental values of the FFCF.[39, 45, 46] Furthermore, a similar result is obtained from the TPZS analysis of the 2DIR spectra of a very diluted sample of azide in D_2O (50 μm), which does not have any appreciable photon echo or interference, but its 2DIR spectra have significant amount of noise (see Supplementary Material). Notably, the TPZS analysis of the diluted azide sample (Figure 3) produces a CS factor with expected time dependence (i.e., exponential decay). More importantly, the derived dynamics for the dilute sample is comparable to the FFCF dynamics obtained from the concentrated sample. However, under these dilute conditions neither CLS nor NLS retrieves a reasonable FFCF dynamics due to the large noise as well as the presence of a water grating signal in the data (Table 2).

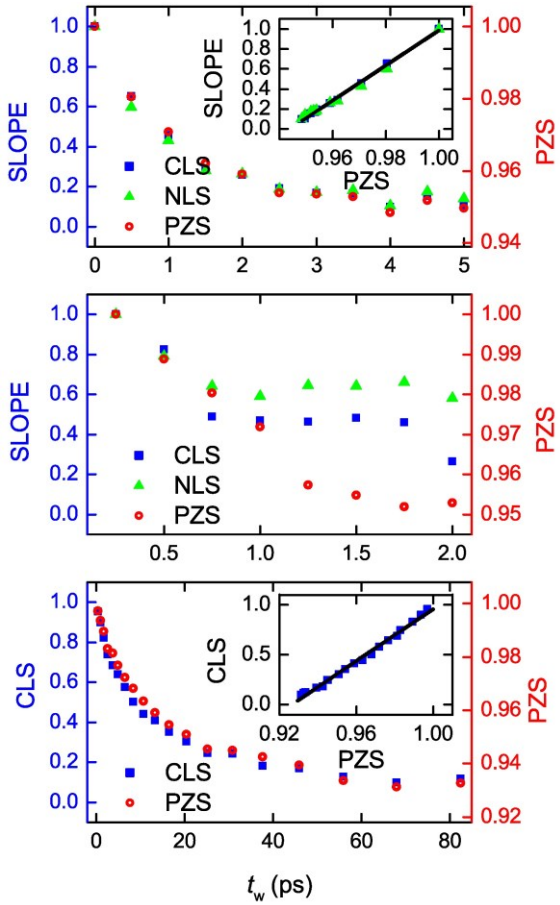

Figure 3. Comparison among FFCF retrieval methods for experimental data. Panels (a) and (b) show analysis of azide samples at 50mM and 50 μM (diluted), respectively, and panel (c) of the

BzSCN in DES using CLS (solid blue square), NLS (solid green triangle) and TPZS (hollow red circle).

Table 3. PZS, CLS and NLS analysis of experimental data. Characteristic time of the dynamics derived from the analysis of simulated 2DIR spectra containing single and two component system.

Sample	τ (ps)									
	TPZS		FR-PZS		CL-PZS		CLS		NLS	
Single transition	ω_1		ω_1		ω_1		ω_1		ω_1	
$N3^-$ in D_2O	1.2 ± 0.1		1.1 ± 0.2		--		1.0 ± 0.4		0.92 ± 0.05	
$N3^-$ in D_2O (dil.)	1.1 ± 0.4		--		--		0.47 ± 0.2		0.25 ± 0.07	
BzSCN in DES	3 ± 1 25 ± 6		3 ± 4 23 ± 23		--		3.4 ± 0.7 22 ± 3		N/A	
Two transitions	ω_1	ω_2	ω_1	ω_2	ω_1	ω_2	ω_1	ω_2	ω_1	ω_2
SCN- in DES	3.1 ± 2.2* 33 ± 8*		1.2 ± 1.1 24 ± 8	12 ± 26 66 ± 208	8.4 ± 8.9 55 ± 78	3.1 ± 1.5 60 ± 16	1.2 ± 0.6* 41 ± 4*		2.2 ± 0.9* 40 ± 14*	
NIPA in D_2O	0.9 ± 0.1	0.8 ± 0.1	--	--	--	--	--		0.96 ± 0.03	N/A

* These results are average values because neither analysis method can separate the two components.

The last sample having a single transition in the investigated region is the BzSCN in a DES. In this case, the nitrile stretch of the BzSCN molecule presents a large anharmonicity, or equivalent a large separation between positive and negative peaks in the 2DIR spectra, and hence, can only be analyzed using the CLS method. In addition, the FFCF dynamics presents a biexponential decay.[41] The TPZS analysis on the BzSCN 2DIR spectra produces similarity factors with a clear time dependence (Figure 3), which in this case is well modeled with a bi-exponential decay function as previously described.[41] The mathematical form of this modeling function is: $y_0 + \sum_i A_i e^{-\frac{t_w}{\tau_{FFCFi}}}$, where y_0 is the minimum PZS similarity factor, and A_i is the amplitude of the i^{th} decay time and τ_{FFCFi} is the i^{th} characteristic time of the decay function. The characteristic times of the decay functions derived from the TPZS analysis are comparable to those derived from using CLS (Table 3). Moreover, as in the previous cases a strong linear correlation is found between the TPZS and CLS metrics indicating that both methods capture the same dynamics, which corroborates the

applicability of the method to vibrational transitions with high anharmonicity. The result is not surprising given that the PZS methodology can also correctly model 2DIR spectra with biexponential dynamics, retrieving both characteristic times and amplitudes, and with similar accuracy as the CLS (see Supplemental Material). Overall, the comparable FFCF dynamics retrieved using either the PZS method or the more traditional methodologies demonstrate that the TPZS method is suitable for extracting the FFCF dynamics of the 2DIR spectra of vibrational modes.

It is demonstrated so far that the PZS correctly assesses the FFCF dynamics from the waiting time dependence of the 2DIR spectra. Notably, the method can also be used to determine the central frequency of the 2DIR peaks giving rise to the observed dynamics. To this end, a frequency resolved PZS (FR-PZS) analysis is required (see methods section). The FR-PZS analysis applied on the simulated 2DIR spectra with a single component (case I and II of Table 1) shows that the τ_{FFCF} as a function of the 2DIR excitation frequency, or equivalent the FR-PZS window frequency, has a minimum (Figure 4). This minimum is located at approximately the frequency position of the transition, which in the case I and II is 2225 cm^{-1} (Table 1). The appearance of a minimum in the FR-PZS plot is not surprising and arises from the fast initial changes with waiting times and lack of change afterwards in the similarity factor at maximum of the peak. In other words, the FR-PZS analysis shows that the region comprising the 2DIR spectra maximum, or equivalent the center of the rotation, observes the changes that rapidly stabilizes as function of waiting time because is the center of the rotation. Furthermore, the characteristic time from the FR-PZS analysis also captures the dynamics of the system as seen on either side of the minimum (Figure 4). However, it appears that the high frequency side produces a more accurate characteristic time. This last result indicates that the FR-PZS also retrieves the FFCF dynamics even when constrained to

small parts of the 2DIR spectra, but the accuracy of the retrieved characteristic times is significantly reduced when the frequency range is reduced for the PZS analysis.

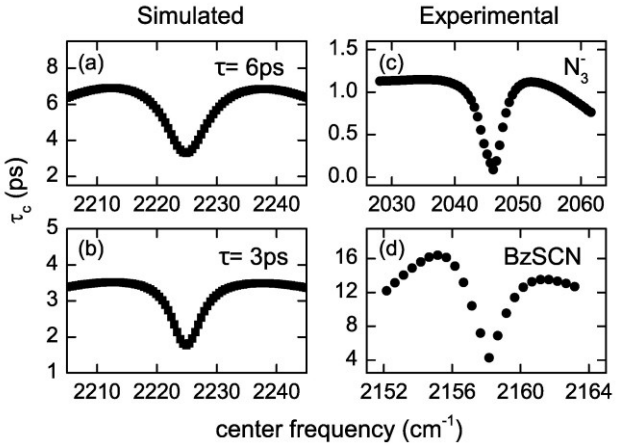

Figure 4. FR-PZS analysis on simulated and experimental data. Panels (a) and (b) show the FFCF characteristic time as a function of the 2DIR excitation frequency for simulated data (case I and I of Table 1), while panels (c) and (d) are for aqueous azide and BzSCN in DES, respectively.

The retrieval of the frequency positions from the FR-PZS analysis of the 2DIR spectra is not limited to theoretical data. The FR-PZS analysis on experimental 2DIR spectra containing a single transition also shows the locations of the peak maximum (Figure 4), which in the case of the azide and BzSCN data corresponds to frequency positions of 2046 cm^{-1} and 2158 cm^{-1}, respectively. These two frequency positions are in very good agreement with the experimental values derived from the FTIR spectra, which are 2043 cm^{-1} for azide ion in water and 2156 cm^{-1} for BzSCN in the DES.[41] Note that the BzSCN sample has a bi-exponential dynamics in the FFCF,[41] and hence, an average of the correlation times was used.[47]

In summary, the results of this section highlight the suitability of the PZS analysis for retrieving the FFCF dynamics from the waiting time evolution of the 2DIR spectra of system having a single vibrational transition irrespective of the anharmonicities and in the presence of a significant amount of noise. In addition, the complementary FR-PZS analysis allow us to retrieve the central frequency position of the vibrational transition giving rise to the observed dynamics. These two

last characteristics of the PZS method provide a unique applicability in congested 2DIR spectra containing multiple transitions.

Multiple transitions

The applicability of the PZS methodology to retrieve the FFCF dynamics and the central frequency position in complex 2DIR spectra containing more than one vibrational transition is shown in this section. The analysis first focused on simulated 2DIR spectra each containing two different vibrational transitions with different frequency separation (unresolved and resolved peaks) and different FFCF dynamics (case III and IV of Table 1).

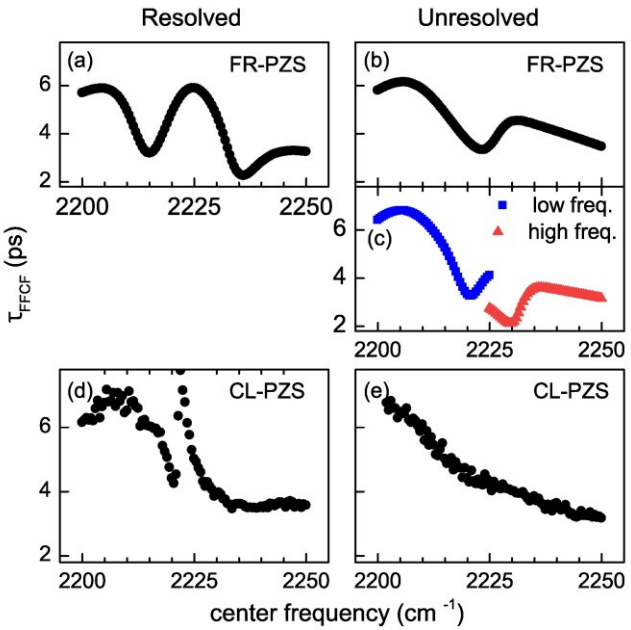

Figure 5. FR-PZS and CL-PZS analysis results for simulated 2DIR spectra with two transitions. Panels (a) and (b) correspond to the FR-PZS analysis for 20 cm^{-1} and 10 cm^{-1} peak separation, respectively. Panel (c) is also FR-PZS analysis but in smaller regions. Panels (d) and (e) correspond to the CL-PZS analysis for 20 cm^{-1} and 10 cm^{-1} peak separation, respectively.

The FR-PZS analysis in 2DIR spectra with spectrally resolved bands (Figure 5) shows the corresponding behavior to the single transition cases (Figure 4), since each band has a local minimum in the FFCF characteristic time as a function of the window frequency. Moreover, the position of the two minima (2215 cm^{-1} and 2236 cm^{-1}) in the frequency resolved characteristic

time agree well with the frequency locations and separation of the peaks used in the simulation (2215 cm^{-1} and 2235 cm^{-1}). In addition, the FR-PZS analysis on the spectra with two resolved transitions also correctly measures the individual characteristic times of the two transitions as seen in the low and high frequency side of the minima. More importantly, these FR-PZS results are in much better agreement with the simulation values than the results produced by either CLS or NLS metrics (Table 3). Finally, the FR-PZS analysis is independent of the dynamics assigned to each transition, as it produces the same result when their FFCF correlation times are reversed or in the presence of FFCFs with multiple decay times (see Supplementary Material).

Notably, in the case of 2DIR spectra with unresolved bands, the frequency resolved characteristic time (Figure 5) shows a function with a single local minimum at the averaged peak frequency in agreement with the observation of a "single" 2DIR peak. However, the frequency resolved profile of the characteristic time is noticeably different from the one obtained for either one (Figure 4) or two resolved transitions (Figure 5). Specifically, there is a mismatch in the characteristic time on either side of the minimum and a slope is observed on the side of the minimum. These features in the profile of the frequency resolved characteristic time are the unequivocal signatures of the presence of more than one underlying transition within the unresolved peak.

The FR-PZS analysis of the 2DIR spectra using smaller excitation frequency windows (see Supplementary Material) shows that the spectra contain two underlying peaks (Figure 5). The FR-PZS analysis locates these transtions approximately at 2222 cm^{-1} and 2232 cm^{-1} with FFCF decay times of 6.8 ps and 3.6 ps, respectively. These results are in very good agreement with the parameters used for simulating the 2DIR spectra. Note that the small errors in the position of the transition frequencies (local minima) and their decay time are likely to arise from the strong overlap between the two vibrational transitions. Moreover, both NLS and CLS capture an averaged

FFCF dynamics when the two vibrational transitions are located very close to one another. It is now clear that the PZS methodology not only demonstrates unambiguously the presence of two transtions even when the two transitions cannot be resolved, but also determines the FFCF dynamics with much lower error when compared to NLS or CLS. However, the FR-PZS methodology is not perfect and retrieves values of the characteristic times, which have errors on the order of 20%.

The redundancy of the PZS methods allows to corroborate the previous finding. For this purpose, the TPZS is applied along the central line (see CL-PZS in the methods section). In this case, the CL-PZS analysis correctly captures the characteristic time of FFCF dynamics for both resolved and unresolved transitions irrespective of the FFCF decay times and their corresponding vibrational transitions. Note that the CL-PZS analysis applied to simulated 2DIR spectra with a single transition also correctly captures its dynamics (see Supplementary Material).

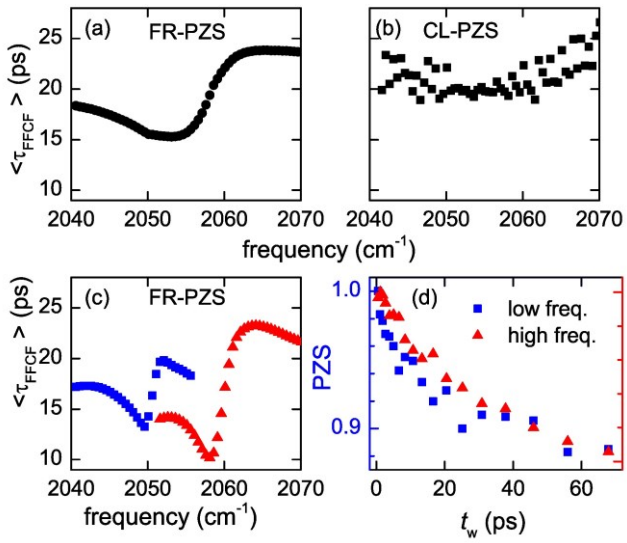

Figure 6. FR-PZS and CL-PZS analysis results for experimental 2DIR spectra with two unresolved transitions. Panels (a) and (b) correspond to characteristic time as a function of the 2DIR excitation frequency derived from the FR-PZS and CL-PZS analysis, respectively, while panel (c) correspond to the FR-PZS analysis in smaller regions. Panel (d) contains the time traces for the low (blue squares) and high (red triangles) frequency transition of the similarity factors derived from the FR-PZS analysis.

Finally, the validity of the PZS methodology in complex 2DIR spectra was demonstrated in experimental data containing two unresolved vibrational transitions. The FR-PZS analysis is first applied to the thiocyanate ion in amide-based DES. This system is particularly complicated because the 2DIR spectra in the nitrile stretch region present two unresolved vibrational transitions that can only be identified from the rhombus-like shape of the 2D peaks at a later waiting time.[43] The FR-PZS analysis on the DES sample (Figure 6) confirms the underlying complexity of the 2DIR spectra as evidenced by the frequency resolved characteristic time does not resemble one with a single transition (Figure 4). Moreover, FR-PZS analysis on sections with smaller excitation frequency ranges reveals the presence and location of two underlying transitions (Figure 6). The analysis also shows that the signatures of the low and high frequency components are observed at 2050 cm^{-1} and 2058 cm^{-1}, respectively. In addition, the modeling of time evolution of the similarity factors derived from FR-PZS shows that each of the two underlying transitions has a bi-exponential dynamics (Figure 6), each with its own time constants (Table 3). Note that the large errors observed in the time constants are associated to the FR-PZS analysis because of the small frequency window used in the spectral region where the signal to noise ratios are likely to be low. Given the resolving power of the FR-PZS analysis, it is not surprising that these newly derived time constants differ significantly from the previously reported CLS modeling.[43] To corroborate the dynamical behavior of the system derived from the FR-PZS, the CL-PZS analysis is performed. This latter analysis corroborates the previous characterizations of the dynamics, as it also shows a faster average dynamic for the low frequency transition when compared to its high frequency counterpart. The reasonable agreement between FR-PZS and CL-PZS analysis further supports the correct determination of the FFCF dynamics from the PZS_{tw} factor using the PZS methodology

and highlights the applicability of the method for complex 2DIR spectra containing overlapping vibrational transitions.

Another example of complex 2DIR spectra correspond to the amide I band of N-isopropylacrylamide (NIPA) in D_2O.[48] In this case, the 2DIR spectra contain two transitions from two vibrational states arising from amide having two different solvation shells that interconvert with ultrafast time scales. The two transitions are fairly well resolved in the spectra, but the appearance and growth of cross peaks with waiting time from the chemical exchange hinders the determination of the FFCF dynamics for the high frequency transition. As in the previous cases, the TPZS analysis on the spectral regions containing each peak (see Supplementary Material) allows us to derive their FFCF dynamics in this complex scenario. The similarity factors derived from TPZS analysis show explicitly a decay in their time evolution (Figure 7). Furthermore, a correlation plot between the CLS and PZS metrics (inset of Figure 7) demonstrate a strong linear correlation between the two metrics (R^2=0.93) indicating that the two methods retrieve essentially the same dynamics. The modeling of the PZS with a single decay function produces similar decay times of 0.9 ± 0.1 ps and 0.8 ± 0.1 ps for time traces corresponding to the 1608 cm^{-1} and 1618 cm^{-1} peaks, respectively. Although it is not possible to evaluate how precise is the value of the FFCF dynamics for the high frequency transition from this analysis, the obtained characteristic times is in agreement with vibrational transition being a NIPPA hydrated state, since most molecules observed FFCF characteristic times close to 1ps in aqueous environments.[48]

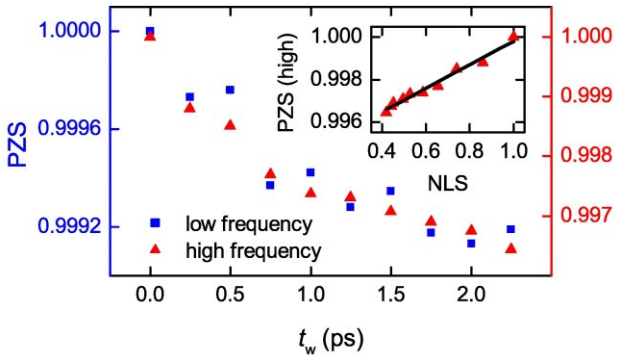

Figure 7. Normalized PZS values as function of waiting times for the low (left) and the high (right) frequency peaks of NIPA 2DIR spectra. Insert shows the correlation plot between the NLS and the PZS for the low frequency transition.

CONCLUSIONS

A new method to analyze 2DIR spectra based on pseudo-Zernike moments (PZM) was developed and studied for retrieving dynamical and spectral information from the 2DIR spectra. The new method is based on using a PZM similarity factor (PZS) to retrieve the information. The practicality of the PZS method was validated with simulated and experimental 2DIR spectra. In the case of a single transition, the PZS method is shown to be capable of obtaining not only the dynamics of spectral diffusion, but also frequency location of the transitions. Notably, the dynamics extracted with this new method are very similar to that obtained by CLS and NLS analysis. In addition, the application of the analysis to more complex 2DIR spectra having two transitions results in the retrieval of the individual dynamics and the center frequency of the transitions even when the two transition result in unresolved spectra. Overall, it is demonstrated in this work that the new PZS analysis method is reliable and can be used in conjunction with standard methods for in-depth analysis of complex 2DIR spectra.

SUPPLEMENTARY MATERIAL

See supplementary material for RMSE plot of simulated 2DIR spectrum as function of PZP order, similarity factors (CS_{tw}, CD_{tw} and PZS_{tw}) as a function of waiting time and their characteristic

decay times, characteristic decay time convergence as a function of n^{th} order PZM, simulated 2DIR spectra of one, CL-PZS analysis on simulated 2DIR spectra with a single transition, PZS vs CLS for simulated 2DIR spectra with biexponential decay, simulated 2DIR spectra with two components with both resolved and unresolved transition as well as single and biexponential decays, FR-PZS and CL-PZS analysis for two component 2DIR spectra with reversed correlation time, FR-PZS analysis for two component 2DIR spectra with biexponential dynamics, 2DIR spectra of all five experimental datasets, and comparison between CLS and TPZS for the amide-based DES 2DIR spectra.

ACKNOWLEDGMENTS

The authors would like to acknowledge financial support from the National Science Foundation (CHE-175135).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Anit Gurung: Data curation (lead); Formal analysis (lead); Investigation (equal); Methodology (equal); Validation (equal); Writing - original draft (supporting); Writing – review & editing (supporting). Daniel G. Kuroda: Conceptualization (lead); Investigation (equal); Methodology (equal); Project administration (lead); Supervision (lead); Validation (supporting); Writing – original draft (equal); Writing – review & editing (lead).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] P. Hamm, and M. Zanni, *Concepts and methods of 2D infrared spectroscopy* (Cambridge University Press, 2011),
[2] D. C. Urbanek *et al.*, J. Phys. Chem. Lett. **1**, 3311 (2010).
[3] K. C. Robben, and C. M. Cheatum, J. Phys. Chem. B **125**, 12876 (2021).
[4] T. Brinzer *et al.*, J. Chem. Phys. **142**, (2015).
[5] K. Kwak, D. E. Rosenfeld, and M. D. Fayer, J. Chem. Phys. **128**, (2008).
[6] K. Kwak *et al.*, J. Chem. Phys. **127**, 124503 (2007).
[7] K. Kwac, and M. H. Cho, J. Chem. Phys. **119**, 2256 (2003).
[8] J. D. Eaves *et al.*, Proc. Natl. Acad. Sci. U.S.A. **102**, 13019 (2005).
[9] S. T. Roberts, J. J. Loparo, and A. Tokmakoff, J. Chem. Phys. **125**, 084502 (2006).
[10] R. Duan *et al.*, J. Chem. Phys. **154**, 174202 (2021).
[11] I. J. Finkelstein *et al.*, Proc. Natl. Acad. Sci. U.S.A. **104**, 2637 (2007).
[12] M. Cho *et al.*, The Journal of Physical Chemistry **100**, 11944 (1996).
[13] W. P. de Boeij, M. S. Pshenichnikov, and D. A. Wiersma, Chem. Phys. Lett. **253**, 53 (1996).
[14] A. Piryatinski, and J. Skinner, J. Phys. Chem. B **106**, 8055 (2002).
[15] C. Fang *et al.*, Proc. Natl. Acad. Sci. U.S.A. **105**, 1472 (2008).
[16] D. Kraemer *et al.*, Proc. Natl. Acad. Sci. U.S.A. **105**, 437 (2008).
[17] J. B. Asbury *et al.*, J. Chem. Phys. **121**, 12431 (2004).
[18] J. B. Asbury *et al.*, J. Phys. Chem. A **108**, 1107 (2004).
[19] E. E. Fenn, and M. Fayer, J. Chem. Phys. **135**, 074502 (2011).
[20] C.-H. Teh, and R. T. Chin, IEEE Trans. Pattern Anal. Mach. Intell. **10**, 496 (1988).
[21] A. Khotanzad, and Y. H. Hong, IEEE Trans. Pattern Anal. Mach. Intell. **12**, 489 (1990).
[22] C.-W. Chong, P. Raveendran, and R. Mukundan, Int. J. Pattern Recognit. Artif. Intell. **17**, 1011 (2003).
[23] L. Li *et al.*, Appl. Opt. **57**, F22 (2018).
[24] H. Gorji, and J. Haddadnia, Neuroscience **305**, 361 (2015).
[25] J. Haddadnia, M. Ahmadi, and K. Faez, EURASIP J. Adv. Signal Process **2003**, 1 (2003).
[26] K. Rahbar, K. Faez, and E. A. Kakhki, J. Opt. Soc. Am. A **30**, 1988 (2013).
[27] M. Bueeler, and M. Mrochen, J Refract Surg **21**, 28 (2005).
[28] P. Allen *et al.*, PLoS One **14**, e0217265 (2019).
[29] v. F. Zernike, Physica **1**, 689 (1934).
[30] C.-W. Chong, P. Raveendran, and R. Mukundan, Pattern Anal. Appl. **6**, 176 (2003).
[31] R. C. Spiers, and J. H. Kalivas, J. Chem. Inf. Model **61**, 2220 (2021).
[32] A. Gurung, and J. H. Kalivas, J. Chemom. **34**, e3245 (2020).
[33] W. Fu, and W. S. Hopkins, J. Phys. Chem. A **122**, 167 (2018).
[34] I. MathWorks, (The MathWorks, Inc. Natick, MA, USA, 2022).
[35] P. Fricker, MATLAB Central File Exchange., 2023).
[36] I. Noda, Appl. Spectrosc. **44**, 550 (1990).
[37] N. H. Ge, M. T. Zanni, and R. M. Hochstrasser, J. Phys. Chem. A **106**, 962 (2002).

[38] A. Ghosh *et al.*, Chem. Phys. Lett. **469**, 325 (2009).
[39] P. Hamm, M. Lim, and R. M. Hochstrasser, Phys. Rev. Lett. **81**, 5326 (1998).
[40] S. Li *et al.*, J. Phys. Chem. B **110**, 18933 (2006).
[41] X. Chen *et al.*, J. Phys. Chem. B **124**, 4762 (2020).
[42] Y. Cui *et al.*, J. Phys. Chem. B **123**, 3984 (2019).
[43] Y. Cui *et al.*, J. Chem. Phys. **155**, 054507 (2021).
[44] Y. Cui, and D. G. Kuroda, J. Phys. Chem. A **122**, 1185 (2018).
[45] X. L. Zhang, R. Kumar, and D. G. Kuroda, J. Chem. Phys. **148**, (2018).
[46] Q. Guo *et al.*, J. Chem. Phys. **142**, (2015).
[47] C. J. Devereux *et al.*, Chem. Phys. **495**, 1 (2017).
[48] E. O. Nachaki, F. M. Leonik, and D. G. Kuroda, J. Phys. Chem. B **126**, 8290 (2022).