

Published in final edited form as:

J Raman Spectrosc. 2016 April ; 47(4): 408–415. doi:10.1002/jrs.4824.

Quantitative, Comparable Coherent Anti-Stokes Raman Scattering (CARS) Spectroscopy: Correcting Errors in Phase Retrieval

Charles H. Camp Jr.^{*}, Young Jong Lee, and Marcus T. Cicerone

Biosystems and Biomaterials Division, National Institute of Standards and Technology, 100 Bureau Dr, Gaithersburg, MD 20899, USA

Abstract

Coherent anti-Stokes Raman scattering (CARS) microspectroscopy has demonstrated significant potential for biological and materials imaging. To date, however, the primary mechanism of disseminating CARS spectroscopic information is through pseudocolor imagery, which explicitly neglects a vast majority of the hyperspectral data. Furthermore, current paradigms in CARS spectral processing do not lend themselves to quantitative sample-to-sample comparability. The primary limitation stems from the need to accurately measure the so-called nonresonant background (NRB) that is used to extract the chemically-sensitive Raman information from the raw spectra. Measurement of the NRB on a pixel-by-pixel basis is a nontrivial task; thus, reference NRB from glass or water are typically utilized, resulting in error between the actual and estimated amplitude and phase. In this manuscript, we present a new methodology for extracting the Raman spectral features that significantly suppresses these errors through phase detrending and scaling. Classic methods of error-correction, such as baseline detrending, are demonstrated to be inaccurate and to simply mask the underlying errors. The theoretical justification is presented by re-developing the theory of phase retrieval via the Kramers-Kronig relation, and we demonstrate that these results are also applicable to maximum entropy method-based phase retrieval. This new error-correction approach is experimentally applied to glycerol spectra and tissue images, demonstrating marked consistency between spectra obtained using different NRB estimates, and between spectra obtained on different instruments. Additionally, in order to facilitate implementation of these approaches, we have made many of the tools described herein available free for download.

Introduction

Coherent anti-Stokes Raman scattering (CARS) is a nonlinear scattering phenomenon in which two photons, “pump” and “Stokes”, coherently excite a molecular vibration. From the excited mode, a “probe” photon inelastically scatters off with an energy increase equal to that of the vibrational state. This optical technique affords label-free, molecularly-sensitive

^{*}To whom correspondances should be address, charles.camp@nist.gov.

Disclaimer

Any mention of commercial products or services is for experimental clarity and does not signify an endorsement or recommendation by the National Institute of Standards and Technology. The authors declare no competing financial interests.

investigation of samples without autofluorescence competition and at significantly higher speeds than offered by traditional spontaneous Raman spectroscopy[1, 2]. CARS microscopies and microspectroscopies have demonstrated success on a large variety of material and biological samples ranging from polymer blends[3, 4] to brain tumor masses [5, 6], requiring fractions of a second for a micrograph highlighting a single vibrational mode to a few minutes for a complete hyperspectral image.

A fundamental challenge to harnessing the information content in CARS microspectroscopies, is extraction of the chemically-specific Raman signal from the so-called “nonresonant background” (NRB). The NRB is predominantly comprised of electronic signal contributions from other nonlinear optical phenomena that are less chemically specific [7]. Although it is sometimes viewed as an interference, the NRB actually amplifies the weak Raman signal, enabling high-sensitivity detection[7, 6]. This coherent mixing, however, does bring spectral distortions. For CARS microscopies that probe small increments of the full vibrational spectrum, physical methods are utilized to reduce the NRB generation, which in turn reduces the overall CARS signal [8, 9, 10, 7]. For spectroscopic CARS techniques (microspectroscopies), however, two classes of numerical methods are commonly used to remove the distortion of the NRB: one based on maximizing entropy [11] and the other utilizing the Kramers-Kronig (KK) relation [12]. These techniques are functionally equivalent [13] and have demonstrated success with materials [14, 15] and biological samples[16, 17, 18, 19, 6]. Both of these techniques, however, rely upon an accurate measurement of the NRB, which also incorporates the system excitation profile and spectral response. To date, no approach has been found to conveniently and accurately measure the NRB, so retrieved spectra generally have phase and amplitude errors. Additionally as typically implemented, these techniques implicitly assume a constancy of the nonresonant background at every pixel across the image, which is typically a poor assumption for heterogeneous samples. The resulting processed spectra contain baseline fluctuations, distorted peaks, and provide limited quantifiable information between molecular species; thus, intra- and inter-sample comparisons are significantly hampered. Additionally, the measurements performed on different CARS instruments— or even the same instrument under altered operating conditions— are not directly comparable.

In this work, we present a new method of processing CARS spectra that suppresses errors resulting from the use of an inexact reference NRB spectra, removing baseline fluctuations and generating spectra that are agnostic to the reference material used. In upcoming sections, we present a theoretical re-examination of the KK relation and errors resultant from reference NRB measurements. Use of the KK presents analytical expressions for correcting these errors, but these results are also broadly applicable to the maximum entropy method (MEM) [11]. Finally, we present simulated and experimental results using neat glycerol and murine pancreas tissues. Spectra are presented from two broadband CARS (BCARS) architectures with dramatically different excitation profiles[19, 20, 6]. The underlying purpose of this manuscript is to enable the extraction of “pre-processed” spectra that are universally comparable in amplitude and shape. This facilitates dissemination of pre-processed spectra as a new “currency” of coherent Raman imaging data.

1 Theory

Classically, the CARS spectral intensity, I_{CARS} , may be described as [21]:

$$I_{CARS}(\omega) \propto \left| \iiint \chi^{(3)}(\omega) E_p(\omega_p) E_s^*(\omega_s) E_{pr}(\omega_{pr}) \times \delta(\omega - \omega_p + \omega_s - \omega_{pr}) d\omega_p d\omega_s d\omega_{pr} \right|^2, \quad (1)$$

where $\chi^{(3)}$ is the nonlinear susceptibility of the sample; E_p , E_s , and E_{pr} are the pump, Stokes, and probe field amplitudes, within the frequency spaces, ω_p , ω_s , and ω_{pr} respectively; and the delta function ensures energy conservation. This equation may be re-written in a more intelligible form[6]:

$$I_{CARS}(\omega) \propto \left| \left\{ \underbrace{[E_s(\omega) \star E_p(\omega)]}_{C_{st}(\omega)} \chi^{(3)} \right\} * E_{pr}(\omega) \right|^2 \quad (2)$$

$$\cong |\tilde{C}_{st}(\omega)|^2 |\tilde{\chi}^{(3)}(\omega)|^2, \quad (3)$$

where \star and $*$ are the cross-correlation and convolution operations, respectively, and C_{st} is the coherent stimulation profile. Equations 1 and 2 are mathematically equivalent. If we assume a spectrally narrow probe source, we can introduce an “effective” stimulation profile, \tilde{C}_{st} , and nonlinear susceptibility, $\tilde{\chi}^{(3)}$, as presented in eq 3, where

$$\tilde{C}_{st}(\omega) \equiv [C_{st}(\omega) * E_{pr}(\omega)] / \int E_{pr}(\omega) d\omega \text{ and } \tilde{\chi}^{(3)}(\omega) \equiv \chi^{(3)}(\omega) * E_{pr}(\omega).$$

The nonlinear susceptibility describes signal contributions from Raman vibrationally-resonant, χ_R , and vibrationally non-resonant, χ_{NR} , sources: $\chi^{(3)}(\omega) = \chi_R(\omega) + \chi_{NR}(\omega)$. To a first degree approximation[22], spontaneous Raman spectra, I_{Raman} , are related to the vibrationally resonant component of the CARS spectra as: $I_{Raman}(\omega) \propto \text{Im}\{\chi_R(\omega)\}$, where ‘Im’ indicates the imaginary component. The purpose of phase retrieval is to ascertain a phase, $\phi(\omega)$, that isolates the Raman resonant components from the total nonlinear susceptibility.

1.1 Phase Retrieval Using the Kramers-Kronig Relation

The KK relation states that there is an explicit relationship between the real and imaginary components of a function, $f(\omega)$, that is causal (analytic[23, 24]); thus, if only the real (or imaginary) part is known, the imaginary (or real) part can be calculated. In CARS and other spectroscopies, neither the real nor imaginary portion of $\chi^{(3)}$ is accessible (*n.b.*: \tilde{C}_{st} in eq 3

is not a causal function). If the function is square-integrable, there also exists an explicit relationship between the complex norm of the function and the phase [25]:

$$\ln |f(\omega)| = -\hat{\mathcal{H}}\{\phi(\omega)\} \quad (4)$$

$$\phi(\omega) = \hat{\mathcal{H}}\{\ln |f(\omega)|\}, \quad (5)$$

where $\hat{\mathcal{H}}$ is the Hilbert transform[26], which assumes knowledge of the complex modulus of the function or phase over an infinite frequency range. Practically, however, the CARS spectral recording window is limited; thus, we will introduce a “windowed” Hilbert transform: $\hat{\mathcal{H}}_W$, as:

$$\hat{\mathcal{H}}_W\{f(x); \omega_a, \omega_b\} = \frac{\mathcal{P}}{\pi} \int_{\omega_a}^{\omega_b} \frac{f(x')}{x - x'} dx' \quad (6)$$

$$\lim_{\substack{\omega_a \rightarrow -\infty \\ \omega_b \rightarrow \infty}} \hat{\mathcal{H}}_W\{f(x); \omega_a, \omega_b\} = \hat{\mathcal{H}}\{f(x)\}, \quad (7)$$

which is limited to the spectral range ω_a to ω_b (for compactness, these parameters will be neglected from the operator form). \mathcal{P} is the Cauchy principle value. Under the conditions that (a) the Raman peaks encompassed within this window are not affected by those outside of the window and (b) any resonances of χ_{NR} are far-removed from those of χ_R (as is typically the case with infrared stimulation), the windowed- and analytic Hilbert transform are related as:

$$\hat{\mathcal{H}}_W\left\{\frac{1}{2}\ln|\tilde{\chi}^{(3)}(\omega)|^2\right\} \approx \hat{\mathcal{H}}\left\{\frac{1}{2}\ln|\tilde{\chi}^{(3)}(\omega)|^2\right\} + \varepsilon(\omega), \quad (8)$$

where $\varepsilon(\omega)$ is an additive error term (see Figure S1 and associated s demonstrating the additive nature). Applying eqs 5 and 8 to eq 3, the retrieved phase from the raw CARS spectrum, ϕ_{CARS} , may be described as:

$$\phi_{CARS}(\omega) = \hat{\mathcal{H}}_W\left\{\frac{1}{2}\ln I_{CARS}(\omega)\right\} \approx \varepsilon(\omega) + \hat{\mathcal{H}}_W\left\{\frac{1}{2}\ln|\tilde{C}_{st}(\omega)|^2\right\} + \hat{\mathcal{H}}\left\{\frac{1}{2}\ln|\tilde{\chi}^{(3)}(\omega)|^2\right\} \quad (9)$$

$$= \varepsilon(\omega) + \hat{\mathcal{H}}_W \left\{ \frac{1}{2} \ln |\tilde{C}_{st}(\omega)|^2 \right\} + \angle [\chi_R(\omega) + \chi_{NR}(\omega)], \quad (10)$$

where \angle denotes the angle (phase). The retrieved phase is not simply that of the nonlinear susceptibility but also contains contributions from the windowing error and the effective stimulation profile. If one can measure the NRB spectrum without Raman components, I_{NRB} , and assuming that the spectrum is far-removed from electronic resonances such that χ_{NR} is approximately real, then the following phase retrieval can be utilized in lieu of eq 10:

$$\begin{aligned} \phi_{CARS/NRB}(\omega) &= \hat{\mathcal{H}}_W \left\{ \frac{1}{2} \ln \frac{I_{CARS}(\omega)}{I_{NRB}(\omega)} \right\} \approx \varepsilon(\omega) \\ &+ \hat{\mathcal{H}}_W \left\{ \frac{1}{2} \ln |\tilde{C}_{st}(\omega)|^2 \right\} \\ &- \left[\varepsilon(\omega) + \hat{\mathcal{H}}_W \left\{ \frac{1}{2} \ln |\tilde{C}_{st}(\omega)|^2 \right\} \right] \\ &+ \angle [\chi_R(\omega) + \chi_{NR}(\omega)] \\ &- \angle_{\chi_{NR}}(\omega) \approx \angle [\chi_R(\omega) + \chi_{NR}(\omega)], \quad (11) \end{aligned}$$

which is analogous to applying the KK relation to I_{CARS}/I_{NRB} . Using this ratio as our signal, the retrieved complex spectra, I_{retr} is:

$$I_{retr}(\omega) = \sqrt{\frac{I_{CARS}(\omega)}{I_{NRB}(\omega)}} \exp i \phi_{CARS/NRB} \approx \frac{|\tilde{\chi}^{(3)}(\omega)|}{|\tilde{\chi}_{NR}(\omega)|} \exp i \angle [\chi_R(\omega) + \chi_{NR}(\omega)], \quad (12)$$

and the Raman-like spectrum, I_{RL} :

$$I_{RL}(\omega) = \text{Im} \{ I_{retr}(\omega) \} \approx \frac{\text{Im} \{ \chi_R(\omega) \}}{|\chi_{NR}|}; \quad (13)$$

thus, the Raman-like spectrum is proportional to the spontaneous Raman spectrum scaled by the nonresonant component. The earliest KK relation results[12] (presenting a very different derivation) utilized $\sqrt{I_{CARS}} \sin \phi_{CARS/NRB}$, which is directly proportional to the spontaneous Raman spectrum, but it implicitly assumes that $\tilde{C}_{st}(\omega)$ is constant and does not account for $\varepsilon(\omega)$. Later work[18], using an identical phase retrieval method (although with a different derivation— see the Supporting Information), identified a need to remove the stimulation profile via I_{CARS}/I_{NRB} . Together these works explain the method of phase retrieval under certain ideal conditions. In the following sections, we will present the ramifications of the real-world condition when the NRB of the sample is not directly measurable. This analysis,

as only enabled by the derivations in eqs 8 through 13, will lead to a direct method to account and correct for these errors.

1.2 Errors from Inaccurate NRB Measurement

Although phase-retrieval is theoretically straightforward, measuring the NRB is technically challenging. To circumvent this limitation, researchers rely on measurements from model materials, such as coverslip glass or water. This practice assumes that the model material does not contain a significant Raman signature and the NRB varies little from material to material. Additionally in practice, any reference material spectral deviations from the actual NRB are assumed to contribute to a slowly-varying baseline that can be subtracted off via various detrending methods. With the increased sensitivity of spectroscopic CARS systems, however, these assumptions appear more and more naive. As demonstrated below, the difference between the NRB and reference measurement does not lead to an additive error, but rather a multiplicative complex error.

We obtain a reference measurement, I_{ref} as a surrogate for a proper NRB measurement. Here, $I_{ref}(\omega) = \xi(\omega)I_{NRB}(\omega)$, and $\xi(\omega)$ is assumed to be real and positive. By applying eq 11:

$$\phi_{CARS/ref}(\omega) = \phi_{CARS/NRB}(\omega) + \underbrace{\mathcal{H}_W \left\{ \frac{1}{2} \ln \frac{1}{\xi(\omega)} \right\}}_{\phi_{err}(\omega)}, \quad (14)$$

we see that the Raman-like spectrum (eq 13), I_{RL-ref} is:

$$I_{RL-ref}(\omega) = \underbrace{\sqrt{\frac{1}{\xi(\omega)}}}_{A_{err}(\omega)} \sqrt{\frac{I_{CARS}(\omega)}{I_{NRB}(\omega)}} \sin [\phi_{CARS/NRB}(\omega) + \phi_{err}(\omega)] \quad (15)$$

From eqs 14 and 15, one can see that the use of a reference measurement leads to both amplitude (A_{err}) and phase (ϕ_{err}) distortions. Accordingly, removing these errors is not simply a matter of subtraction. The phase error, however, is additive in nature, and connected to the amplitude error via the KK relation:

$$\phi_{err}(\omega) = \mathcal{H} \{ \ln A_{err}(\omega) \} \quad (16)$$

$$\ln A_{err}(\omega) = - \mathcal{H} \{ \phi_{err}(\omega) \}. \quad (17)$$

There is, however, an ambiguity in this relationship. If $\xi(\omega)$ is multiplied by a constant, $\Xi: \phi_{err}(\omega) = \hat{\mathcal{H}} \{ \ln 1 / \Xi(\omega) \} = \hat{\mathcal{H}} \{ \ln 1 / \xi(\omega) \}$, since the Hilbert transform of a constant is zero.

1.3 Correcting Phase Error and Scale

The purpose of correcting for Raman signature extraction errors is not simply to generate qualitatively accurate spectra, but those that are quantitatively reliable, facilitating direct comparison and analysis of spectra collected of different samples with potentially different reference materials, and on various spectroscopic systems having different excitation profiles. In the previous subsection we demonstrated that the use of a reference NRB which only approximates the nonresonant response of the material induces amplitude and phase distortions. Additionally, the commonly used method of subtracting baseline fluctuations, borrowed from the spontaneous Raman community, does not actually remove the errors as (1) the nature of the error is complex valued and (2) the amplitude error is multiplicative.

Below we show that one can properly correct for signal extraction error using the following steps:

1. Remove phase error via phase detrending, and correct for part of the amplitude error via the KK relation
2. Correct for scaling error (involving Ξ) and the error from the windowed Hilbert transform (of ϕ_{err}) via unity-centering of the real component of the retrieved (phase-corrected) spectrum.

As displayed in eq 14, the difference between the ideal phase retrieval (in which the NRB of the sample is exactly known) and that using a model material is ϕ_{err} , which is additive. The retrieved phase (ideal) is qualitative similar to Raman-like spectra in that the spectral features are peaks that extend positively from a zero baseline. A slowly-varying phase error will cause a slowly-varying deviation from the zero baseline. Finding ϕ_{err} ; therefore, is a matter of isolating the erroneous baseline, for which one may use traditional baseline detrending methods (albeit, applied to the phase rather than the Raman-like spectral amplitude). From this extracted ϕ_{err} , using eq 17, one can find the amplitude error. With these variables in hand, one can multiply the retrieved complex spectrum by a complex phase-correction multiplier, generating a phase-corrected spectrum, I_{pc} :

$$I_{pc}(\omega) = \sqrt{\frac{I_{CARS}(\omega)}{I_{ref}(\omega)}} \exp \left[i \phi_{CARS/ref}(\omega) \right] \left\{ \frac{1}{\exp \left[-\hat{\mathcal{H}} \{ \phi_{err}(\omega) \} \right]} \exp \left[-i \phi_{err}(\omega) \right] \right\} \quad (18)$$

n.b.: the use of “phase-corrected” in this context is unrelated to that in Ref. 18. As described in the previous subsection, however, there is an ambiguity in the KK relation between the error phase and amplitude. Thus the calculated A_{err} is only accurate to within a constant multiplier. Additionally, the Hilbert transform in eq 18 is actually a windowed Hilbert transform; thus, $\hat{\mathcal{H}} \{ \phi_{err}(\omega) \} = \hat{\mathcal{H}}_W \{ \phi_{err}(\omega) \} + \varepsilon_{err}(\omega)$, where ε_{err} is a window-effect error term similar to that introduced in eq 8.

To finalize the error correction, one needs to account for the A_{err} ambiguity and ϵ_{err} . Both of these quantities are discoverable by analyzing the real component of the phase-corrected spectrum in eq 18 since the real component of eq 12 is unity-centered, i.e.,

$\left\langle \left| \tilde{\chi}^{(3)} \right| / \left| \tilde{\chi}_{NR} \right| \cos \phi_{CARS/NRB} \right\rangle = 1$. The existence of Ξ , however, will alter the mean; thus, one could calculate the mean of the real component of the retrieved spectrum and normalize the real and imaginary components by this value. ϵ_{err} , however, may impart a frequency-dependent component to this mean. Using numerical means, though, one can find a slowly-varying center-line and normalize the phase-corrected spectrum; thus, removing Ξ and ϵ_{err} in one step. Assuming this centerline can be found, a complex corrected spectrum $I_{corrected}$ may be calculated:

$$I_{corrected}(\omega) = \frac{I_{pr}(\omega)}{\langle \text{Re} \{ I_{pr}(\omega) \} \rangle(\omega)} \quad (19)$$

$$= \frac{\left| \tilde{\chi}^{(3)}(\omega) \right|}{\left| \tilde{\chi}_{NR}(\omega) \right|} \exp i \phi_{CARS/NRB}(\omega), \quad (20)$$

where we have noted the frequency-dependence of the mean-line in the denominator of eq 19. Comparison of eqs 12 and 20 shows that using the prescribed steps, one can retrieve the same spectrum using a reference NRB as if the NRB were measurable.

2 Materials and Methods

2.1 CARS Microspectroscopy

In this manuscript, most CARS spectra were collected on a recently introduced broadband CARS (BCARS) system that has been described elsewhere[6]. In order to demonstrate that properly retrieved spectra can be essentially identical, irrespective of instrumentation, some spectra were collected on an instrument that uses an earlier implementation of BCARS signal generation [19, 14]. The newer BCARS system excites molecular vibrations more efficiently with the highest response at the lowest wavenumbers, whereas, the traditional BCARS system excites most Raman transitions with relatively uniform response. For the newer BCARS system, the total average incident power was <25 mW (3.5 ms integration time), and for the traditional BCARS system <60 mW (7.8 ms integration time).

2.2 CARS Simulations

The CARS simulation software was developed in MATLAB (Mathworks) and numerically implements eq 2 directly. All excitation sources were simulated as real Gaussian functions, and the Raman response a complex Lorentzian (damped harmonic oscillator) as:

$$\chi_R(\omega) = \sum_m \frac{A_m}{\Omega_m - \omega - i\Gamma_m}, \quad (21)$$

where A_m , Ω_m , and Γ_m describe the amplitude [multiplier], wavenumber, and half-width of the m^{th} Raman peak. Specific parameters for the simulations are provided below.

2.3 Signal Pre-Processing

Image and spectral processing are performed using in-house software developed in MATLAB. Specific details of the pre-processing of BCARS spectra (or images) are presented in the Supporting Information (see Figures S2 – S11 and Table S1). In brief, dark spectra are collected as are NRB spectra from reference materials (glass coverslip, glass microscope slide, or water). For noise reduction of BCARS hyperspectral data, we utilize singular value decomposition (SVD)[27] on Anscombe transformed spectra. The Anscombe transform[28, 29] normalizes the noise variance, accounting for mixed Poisson-Gaussian noise. Pertinent singular values are selected by noise analysis in the spectral and spatial domains in an automated or semi-automated fashion (Figures S8, S9). Once SVD is performed, the variance-stabilized, noise-reduced spectra are returned to their normal mixed-noise state using an optimized, generalized inverse Anscombe transformation [29].

Raman-like spectra are retrieved using the Hilbert transform implementation of the KK relation (described later). The erroneous component of the retrieved phase is found in an automated fashion using an asymmetric least-squares (ALS) technique with a Whittaker smoother[30, 31, 32]. Phase and partial amplitude correction are performed as described in eq 18. To determine the mean trend line for final spectral correction (eq 19), a Savitzky-Golay filter is utilized.

2.3.1 Phase Retrieval Using the Hilbert Transform—The Hilbert transform (eq 6) is implemented in the time-domain (t):

$$\hat{\mathcal{H}}_w \{f(\omega)\} = \mathcal{F} \left\{ i \operatorname{sgn}(t) \mathcal{F}^{-1} \{f(\omega)\} \right\}, \quad (22)$$

where \mathcal{F} and \mathcal{F}^{-1} are the Fourier and inverse Fourier transforms, respectively, $\operatorname{sgn}(t)$ is the signum (“sign”) function and $f(\omega)$ is a spectrally dependent function (e.g., $I_{\text{CARS}}/I_{\text{ref}}$). Additionally, we implement a spectral padding procedure [11] to extend the window range, reducing numerical edge effects. This method efficiently retrieves phase with only two Fourier transforms and was designed for parallel processing. 100 parallel solutions, for example, with each spectrum containing 1000 spectral points requires $\sim 200 \mu\text{s}$ per spectrum on a personal computer (16 GB RAM, 3.4 GHz quad-core processor). The KK and associated Hilbert transform code is freely available at <http://github.com/CoherentRamanNIST> in the MATLAB and Python languages.

3 Results

3.1 Simulated Spectra

To validate the presented theory on phase retrieval and error correction, we begin with the simplified case of a two-peak Raman system with parameters (eq 21): $A_1 = 0.25$, $\Omega_1 = 1000 \text{ cm}^{-1}$, $\Gamma_1 = 10 \text{ cm}^{-1}$, $A_2 = 1$, $\Omega_2 = 3100 \text{ cm}^{-1}$, and $\Gamma_2 = 20 \text{ cm}^{-1}$. $\chi_{NR} = 0.55$ and χ_{ref} is χ_{NR} multiplied by a Gaussian function. The simulated nonlinear susceptibilities are presented in Figure 1 a. The CARS spectra that result from these susceptibilities are shown in Figure 1 b.

Figure 2 a shows the retrieved spectra using a reference or the actual NRB ("ideal"). The reference-retrieved spectrum shows a clear, large baseline and distorted (asymmetric) peak shapes. Additionally, the peak amplitudes are sufficiently perturbed that the 1000 cm^{-1} peak appears ~50% larger than the 3100 cm^{-1} , although the latter should be the larger of the two. Figure 2 b shows the difference () between the ideal and nonideal retrieved spectra. From this, one can gather that the traditional tactic of baseline detrending will only resolve the slowly-varying baseline, but the underlying peak-errors (amplitude and phase) will remain. Figure 2 c shows the phase retrieved under ideal and nonideal conditions, which does not display any obvious peak distortions. As clearly presented from the difference of the retrieved phases (Figure 2 d), there is no spectral distortion of the Raman peaks, but only the slowly-varying baseline (ϕ_{err} in eq 14).

Using this phase error and applying a calculated amplitude correction (eq 18), the baseline and asymmetric spectral distortions are removed entirely, as shown in Figure 3 a (differences shown in Figure 3 b). For reference, a traditional amplitude detrending is also displayed, showing the remaining distortions clearly. Although the phase-corrected spectrum is qualitatively similar to the ideal, the relative amplitudes of the two peaks are still incorrect, owing to e_{err} and the amplitude ambiguity described previously. Figure 3 c shows the real part of the phase-corrected spectrum and the mean trend-line that deviates from unity. Using this trend as a scaling factor (eq 19), Figure 3 d demonstrates that the phase-corrected spectrum is now identical to the ideal retrieval (difference shown in Figure 3 e).

As presented elsewhere[13], the KK and MEM phase retrieval methods are functionally equivalent. Figure S12 demonstrates the applicability of the presented phase error correction method to Raman-like spectra extracted from simulated BCARS spectra via the MEM method. Like the KK demonstration above, the prescribed method enables a reference NRB spectrum to be utilized and generates a corrected Raman-like spectrum that is equivalent to one extracted when the NRB is exactly known.

3.2 Experimentally Measured Spectra

The developed error correction method can readily be applied to experimental results without modification. Additionally, this method provides spectra that are comparable between microscopy platforms. Figure S13 a shows CARS spectra collected for neat glycerol on two BCARS systems (average of 100 acquisitions), demonstrating widely different system responses for the same molecule. Figure S13 b shows the recorded reference spectra for 3 different commonly used model materials. These reference spectra

not only demonstrate a great variety of overall amplitude but also spectral features. As expected, retrieving the Raman-like spectra using these references produces amplitude and phase errors, resulting in distortions as shown in Figure S13 c. Figure 4 a shows the Raman-like spectra with the slowly-varying baseline removed. The spectrum retrieved using water demonstrates the most severe distortions. Even the spectra using glasses (slide and coverslip) demonstrate differences. In comparison, Figure 4 b shows the same four spectra after full correction (eq 19). The four spectra are significantly more similar in amplitude and shape. Additionally, one should notice the partial recovery of the OH-stretch peaks ($\sim 3300\text{ cm}^{-1}$) for the spectrum retrieved using water. When a particular reference material is utilized, the retrieved spectra will have suppressed peaks wherever the model material contains Raman peaks. Within spectroscopic CARS literature, coverslip or slide glass have often served as a convenient reference. What was not apparent at the time, however, is that these glasses have nontrivial, glass-dependent Raman peaks. The primary cause of the spectral differences in Figure 4 b are due to these reference material Raman peaks. Figure S15 b shows the retrieved (and corrected) spectra of these reference materials, with their peaks exactly correlating with the deviations in the retrieved spectra (Figure S15 a). These spectra were collected using a time-windowing, self-referencing method (see Supporting Information and Figure S14). This enables collecting an NRB-dominant spectrum directly from the sample, which can then be used as the reference. This also enables the use of reference material spectra with their Raman peaks suppressed. Figure S15 c shows Raman-like (corrected) spectra of glycerol using different references with their peaks suppressed.

3.3 Tissue Imaging

The presented phase retrieval method can also be reliably applied to hyperspectral images. For this purpose, we imaged a histological section of murine pancreatic artery. A $200\text{ }\mu\text{m} \times 200\text{ }\mu\text{m}$ section (90,000 pixels total) was imaged with 3.5 ms dwell times. Reference spectra were collected from water and the sample coverslip. The raw BCARS image was de-noised using SVD on Anscombe-stabilized spectra, keeping 23 singular values. After this de-noising, the hyperspectral data were processed four times: twice with each reference spectrum and twice with amplitude or phase detrending methods. Figures 5 (a,c) are the pseudocolor images of the murine skin highlighting protein in blue (2937 cm^{-1} – 2882 cm^{-1}), DNA in orange (785 cm^{-1}), and in red a shoulder peak that is dominant in the smooth muscle (1339 cm^{-1}) and is tentatively assigned to actin/myosin[33, 34]. The Supporting Information provides more detail regarding how the images were processed. The left-halves of Figures 5 (a,c) show the processing performed using the coverslip NRB reference, and the right-halves using water. The color intensity range of red, green, and blue channels are the same for both half-images within a single figure; although, the range is different between Figure 5 a and Figure 5 c. With amplitude detrending (Figure 5 a), the boundary between the half-images is obvious. Using the coverslip reference, the blue channel (protein) is more intense than when using water. Conversely, the red channel (actin/myosin) is suppressed when using coverslip. The DNA/RNA signature is nearly the same in both images. This highlights that distortions vary across the spectrum, and one cannot simply normalize out the errors. Using phase detrending (and scaling) in Figure 5 b, there is no obvious discontinuity between the two half-images. Figure 5b shows single-pixel spectra collected from a portion of the internal elastic lamina (marked by a white 'x' in Figures 5 a) and corrected using

amplitude detrending. In these spectra the coverslip-processed spectrum is ~40% stronger at higher energies, but ~15% to 30% weaker within the fingerprint region. With phase detrending, shown in Figure 5 d, the peak differences are significantly less obvious, with the predominant cause of residual error being from Raman peaks inherent to the reference materials (in Figures 5 (b,c) significant perturbations induced by reference NRB Raman peaks in coverslip glass are denoted with a '*' and dashed lines). Figure S16 shows a histogram comparison of the Raman peak amplitudes whether performing amplitude detrending or phase detrending (and scaling), demonstrating significantly increased similarity when using the developed method. To quantify the improvement, we calculated on a pixel-by-pixel basis the relative difference (Figure 5 e) between the peaks used to construct Figures 5 (a,c), i.e., the coverslip-processed intensity minus the water-processed intensity, over the mean. Using amplitude detrending, the mean relative difference is 10.5% (standard deviation, σ , 3.7%) for protein, -17.4% (σ =8.4%) for actin/myosin, and -1.1% (σ = 0.5%) for DNA/RNA. Using phase detrending and scaling, the mean relative difference for protein improves to 3.5% (σ = 4.5%) and to 0.08% (σ = 1.1%) for actin/myosin. The mean relative difference for DNA/RNA remains nearly the same (in amplitude) at 1.7% (σ = 0.5%).

4 Summary and Discussion

The aim of this work is to foster quantitative reliability and repeatability of CARS spectra and to promote quantitative analysis of hyperspectral data cubes rather than rely on the pseudocolor imagery. Although this work presents the first efforts towards making CARS spectra reliable, repeatable, and universally comparable, there are still many avenues to improve upon this method. For example, the retrieved spectra using this (and other) methods are normalized to the NRB. As described in eq 12, this results from the practical necessity of removing the system response, which may contain spectral fluctuations due to laser source spectral profiles or filter set characteristics. One solution may be a thoroughly characterized reference material of which the resonant and nonresonant nonlinear susceptibilities are known; thus, enabling the extraction of the system response. This extracted system response could then be applied to the analyses of future CARS acquisitions on that system. Another opportunity is the examination of the proposed method under the condition that the nonresonant nonlinear susceptibility is not approximately real, but rather complex. The ramifications of a complex NRB may become more appreciable due to water absorption as CARS systems move further into the infrared or due to multiphoton electronic absorption as more exotic fluorescent species such as plants or algal cells are investigated.

In conclusion, this work presents an expanded method of Raman signal extraction that corrects for amplitude and phase errors that are ubiquitous in CARS microspectroscopy. The presented technique, does not require any augmentation to current acquisition work flows and performs the corrections *in silico*. The corrected spectra, as demonstrated with neat liquids and tissues images, significantly reduces the intra-spectral distortions caused by the use of NRB reference spectra, and facilitates direct, quantitative comparison between samples and microscopy systems. In the future, it is the hope of the authors that this method (and future improvements) will enable mass dissemination of coherent Raman hyperspectral data cubes for community data mining and analysis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Zumbusch A, Holtom G, Xie XS. Three-dimensional vibrational imaging by coherent anti-Stokes Raman scattering. *Phys Rev Lett*. 1999; 82:4142–4145.
2. Petrov GI, Arora R, Yakovlev VV, Wang X, Sokolov AV, Scully MO. Comparison of coherent and spontaneous Raman microspectroscopies for noninvasive detection of single bacterial endospores. *Proc Natl Acad Sci U S A*. 2007; 104:7776–7779. [PubMed: 17483468]
3. Kee TW, Cicerone MT. Simple approach to one-laser, broadband coherent anti-Stokes Raman scattering microscopy. *Opt Lett*. 2004; 29:2701–2703. [PubMed: 15605477]
4. von Vacano B, Meyer L, Motzkus M. Rapid polymer blend imaging with quantitative broadband multiplex CARS microscopy. *J Raman Spectrosc*. 2007; 38:916–926.
5. Evans CL, Xu X, Kesari S, Xie XS, Wong STC, Young GS. Chemically-selective imaging of brain structures with CARS microscopy. *Opt Express*. 2007; 15:12076–12087. [PubMed: 19547572]
6. Camp CH Jr, Lee YJ, Heddleston JM, Hartshorn CM, Walker ARH, Rich JN, Lathia JD, Cicerone MT. High-speed coherent Raman fingerprint imaging of biological tissues. *Nat Photonics*. 2014; 8:627–634. [PubMed: 25621002]
7. Müller M, Zumbusch A. Coherent anti-Stokes Raman Scattering Microscopy. *ChemPhysChem*. 2007; 8:2156–2170. [PubMed: 17768730]
8. Cheng J-X, Book LD, Xie XS. Polarization coherent anti-Stokes Raman scattering microscopy. *Opt Lett*. 2001; 26:1341–1343. [PubMed: 18049602]
9. Ganikhanov F, Evans CL, Saar BG, Xie XS. High-sensitivity vibrational imaging with frequency modulation coherent anti-Stokes Raman scattering (FM CARS) microscopy. *Opt Lett*. 2006; 31:1872–1874. [PubMed: 16729099]
10. Dudovich N, Oron D, Silberberg Y. Single-pulse coherently controlled nonlinear Raman spectroscopy and microscopy. *Nature*. 2002; 418:512–514. [PubMed: 12152073]
11. Vartiainen EM. Phase retrieval approach for coherent anti-Stokes Raman scattering spectrum analysis. *J Opt Soc Am B*. 1992; 9:1209–1214.
12. Liu Y, Lee YJ, Cicerone MT. Broadband CARS spectral phase retrieval using a time-domain Kramers-Kronig transform. *Opt Lett*. 2009; 34:1363–1365. [PubMed: 19412273]
13. Cicerone MT, Aamer KA, Lee YJ, Vartiainen EM. Maximum entropy and time-domain Kramers-Kronig phase retrieval approaches are functionally equivalent for CARS microspectroscopy. *J Raman Spectrosc*. 2012; 43:637–643.
14. Lee YJ, Moon D, Migler KB, Cicerone MT. Quantitative image analysis of broadband CARS hyperspectral images of polymer blends. *Anal Chem*. 2011; 83:2733–2739. [PubMed: 21395296]
15. Baldacchini T, Zimmerley M, Kuo C-H, Potma EO, Zadayan R. Characterization of microstructures fabricated by two-photon polymerization using coherent anti-stokes Raman scattering microscopy. *J Phys Chem B*. 2009; 113:12663–12668. [PubMed: 19715350]
16. Rinia HA, Bonn M, Müller M, Vartiainen EM. Quantitative CARS spectroscopy using the maximum entropy method: the main lipid phase transition. *Chemphyschem*. 2007; 8:279–287. [PubMed: 17177224]
17. Rinia HA, Bonn M, Vartiainen EM, Schaffer CB, Müller M. Spectroscopic analysis of the oxygenation state of hemoglobin using coherent anti-Stokes Raman scattering. *J Biomed Opt*. 2006; 11:050502. [PubMed: 17092146]
18. Masia F, Glen A, Stephens P, Borri P, Langbein W. Quantitative chemical imaging and unsupervised analysis using hyperspectral coherent anti-Stokes Raman scattering microscopy. *Anal Chem*. 2013; 85:10820–10828. [PubMed: 24099603]
19. Parekh SH, Lee YJ, Aamer KA, Cicerone MT. Label-free cellular imaging by broadband coherent anti-Stokes Raman scattering microscopy. *Biophys J*. 2010; 99:2695–2704. [PubMed: 20959111]

20. Lee YJ, Parekh SH, Kim YH, Cicerone MT. Optimized continuum from a photonic crystal fiber for broadband time-resolved coherent anti-Stokes Raman scattering. *Opt Express*. 2010; 18:4371–4379. [PubMed: 20389449]
21. Gomez, JS. Coherent Raman Spectroscopy. In: Laserna, JJ., editor. *Mod Tech Raman Spectrosc*. John Wiley & Sons; Chichester: 1996. p. 305-342.chap 8
22. Tolles WM, Nibler JW, McDonald JR, Harvey AB. A review of the theory and application of coherent anti-Stokes Raman spectroscopy (CARS). *Appl Spectrosc*. 1977; 31:253–271.
23. Toll JS. Causality and the Dispersion Relation: Logical Foundations. *Phys Rev*. 1956; 104:1760–1770.
24. Lucarini, V., Saarinen, JJ., Peiponen, K-E., Vartiainen, EM. *Kramers-Kronig Relations in Optical Materials Research*. Springer-Verlag; Berlin, Germany: 2005.
25. Smith DY. Dispersion relations for complex reflectivities. *J Opt Soc Am*. 1977; 67:570–571.
26. Poularikas, AD. The Hilbert Transform. In: Dorf, RC., editor. *Handbook Formulas Tables Signal Processing*. CRC; Boca Raton: 1999. chap. 15
27. Cicerone, MT., Lee, YJ., Parekh, SH., Aamer, KA. Photonic crystal fiber-based broadband CARS microscopy. In: Cheng, J-X., Xie, XS., editors. *Coherent Raman Scattering Microscopy*. CRC; Boca Raton, FL: 2012. p. 329-352.chap 15
28. Seymour, L., Starck, J-L., Murtagh, FD., Bijaoui, A. *The Multiscale Approach*. Cambridge University Press; Cambridge: 1998. *Image Processing and Data Analysis*.
29. Mäkitalo M, Foi A. Optimal inversion of the generalized Anscombe transformation for Poisson-Gaussian noise. *IEEE Trans Image Process*. 2013; 22:91–103. [PubMed: 22692910]
30. Eilers PHC, Boelens HFM. Baseline Correction with Asymmetric Least Squares Smoothing. 2005 Unpublished.
31. Eilers PHC. A perfect smoother. *Anal Chem*. 2003; 75:3631–3636. [PubMed: 14570219]
32. Urbas AA, Choquette SJ. Automated spectral smoothing with spatially adaptive penalized least squares. *Appl Spectrosc*. 2011; 65:665–677. [PubMed: 21639989]
33. Romer TJ, Brennan JF, Fitzmaurice M, Feldstein ML, Deinum G, Myles JL, Kramer JR, Lees RS, Feld MS. Histopathology of human coronary atherosclerosis by quantifying its chemical composition with Raman spectroscopy. *Circulation*. 1998; 97:878–885. [PubMed: 9521336]
34. Buschman HP, Deinum G, Motz JT, Fitzmaurice M, Kramer JR, Van Der Laarse A, Bruschke AV, Feld MS. Raman microspectroscopy of human coronary atherosclerosis: Biochemical assessment of cellular and extracellular morphologic structures in situ. *Cardiovasc Pathol*. 2001; 10:69–82. [PubMed: 11425600]

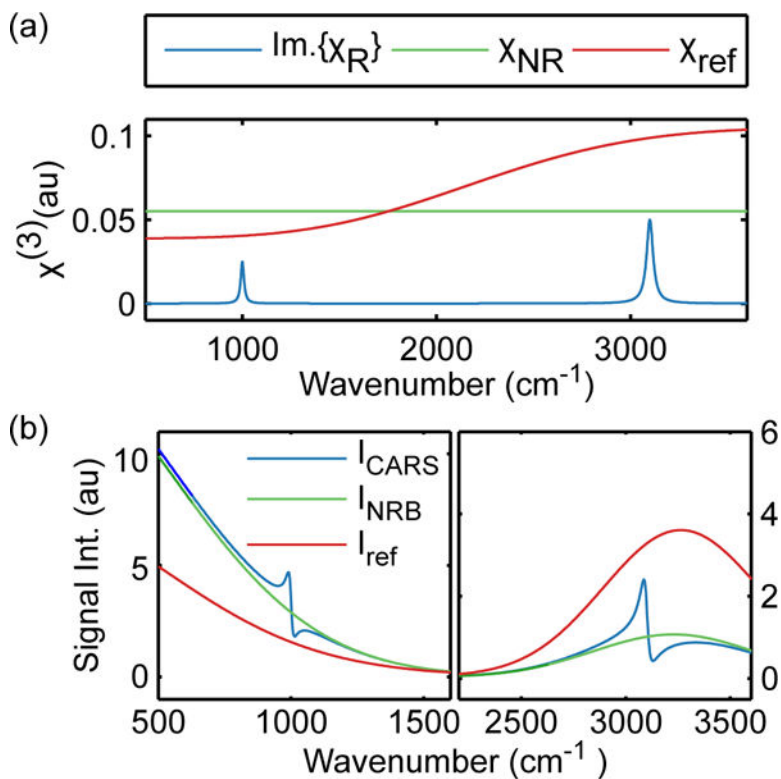


Figure 1.

Simulated nonlinear susceptibilities and CARS spectra. (a) Raman (χ_R) and nonresonant nonlinear susceptibilities of the sample (χ_{NR}) and of a reference (χ_{ref}). (b) Simulated CARS spectra of the sample (I_{CARS} , $\chi = \chi_R + \chi_{NR}$), the sample in absence of Raman components (I_{NRB}), and of the reference material (I_{ref}).

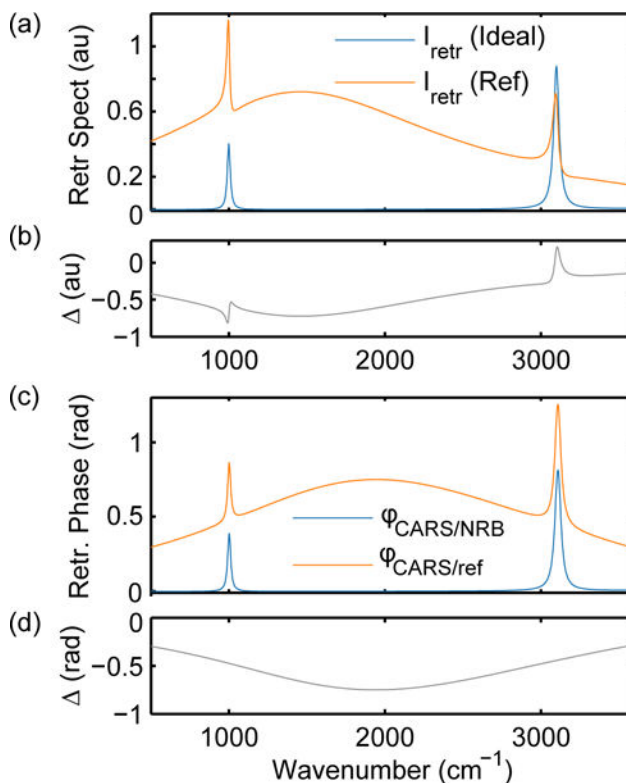


Figure 2.

Comparison of retrieved spectra using the KK under ideal conditions and with use of an NRB reference. (a) Retrieved spectrum (I_{retr}) when the NRB is exactly known (Ideal) and when a reference is utilized (Ref). (b) The difference, Δ , between the ideal and reference retrieval showing remnants of the original Raman peaks. (c) The KK-retrieved phase when the NRB is known ($\phi_{\text{CARS/NRB}}$) and when using a reference ($\phi_{\text{CARS/ref}}$). The phase difference, Δ , between the ideal and reference retrieval that shows only a smooth baseline with no Raman peak information.

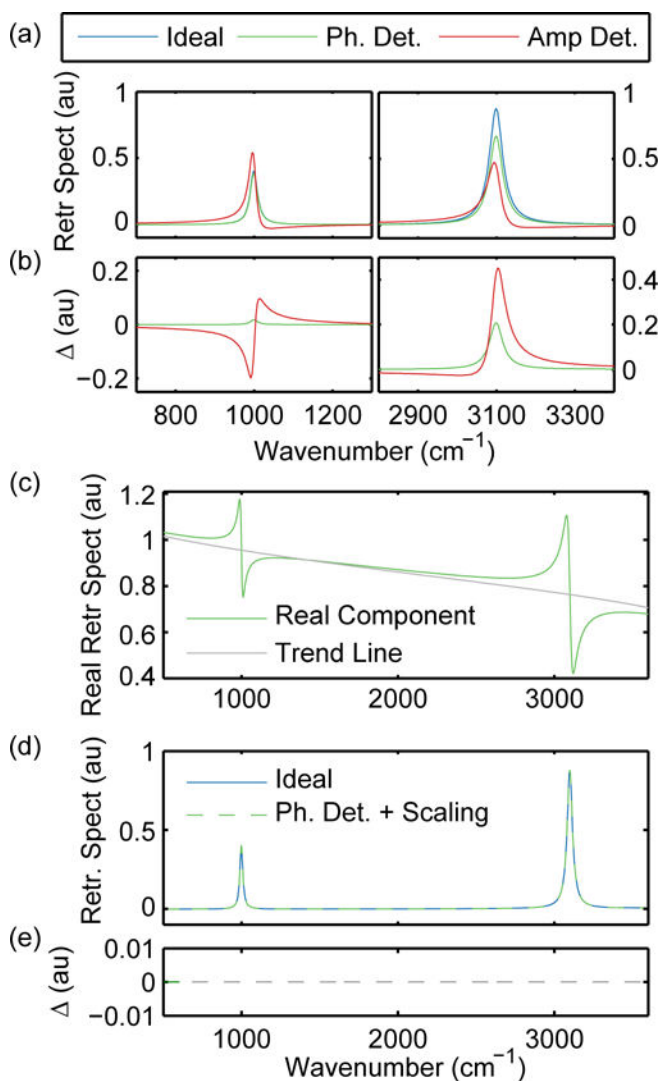


Figure 3.

Residual error and their correction using the proposed phase-detrending and scaling with comparison to tradition amplitude detrending. (a) KK-retrieved spectra when the NRB is known (Ideal) and using a reference material with phase detrending (Ph. Det., see eq 18) and amplitude detrending (Amp. Det.). (b) The difference, Δ , between the ideal and corrected spectra. Amplitude detrending creates asymmetric peak distortion not present with phase detrending. (c) Using the mean-trend of the real component of the retrieved spectrum allows for correction of edge effects and scaling ambiguity. (d) Phase-detrended and scaled spectrum (Ph. Det. + Scaling, imaginary portion of eq 19) is identical to the ideal retrieval (Ideal), with (e) no difference, $\Delta = 0$.

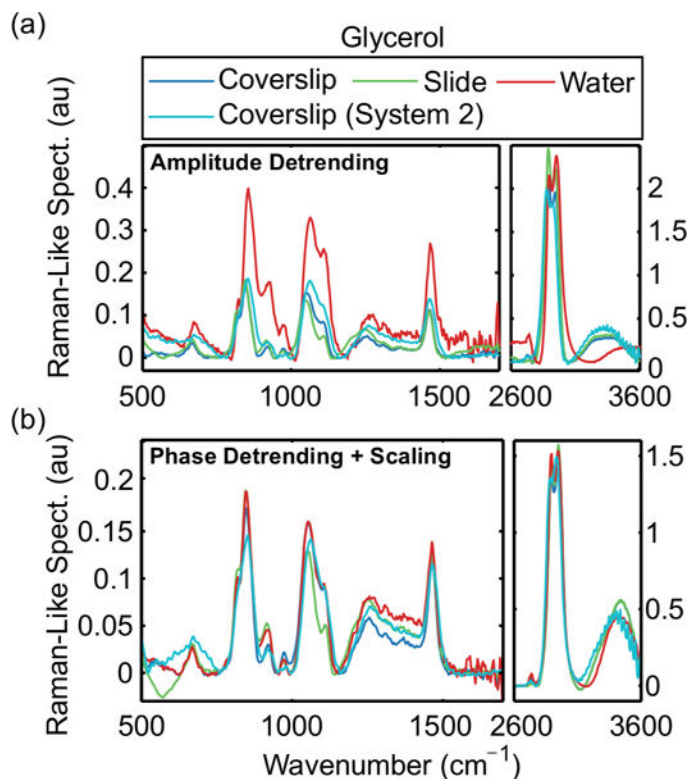


Figure 4.

Spectral correction of neat glycerol. (a) KK-retrieved Raman-like spectra using NRB reference materials with traditional amplitude-detrending, highlighting the inability of baseline detrending to remove amplitude and phase error. (b) Corrected spectra with phase detrending and scaling (imaginary portion of eq 19), showing close agreement with residual differences primarily arising from reference material Raman peaks.

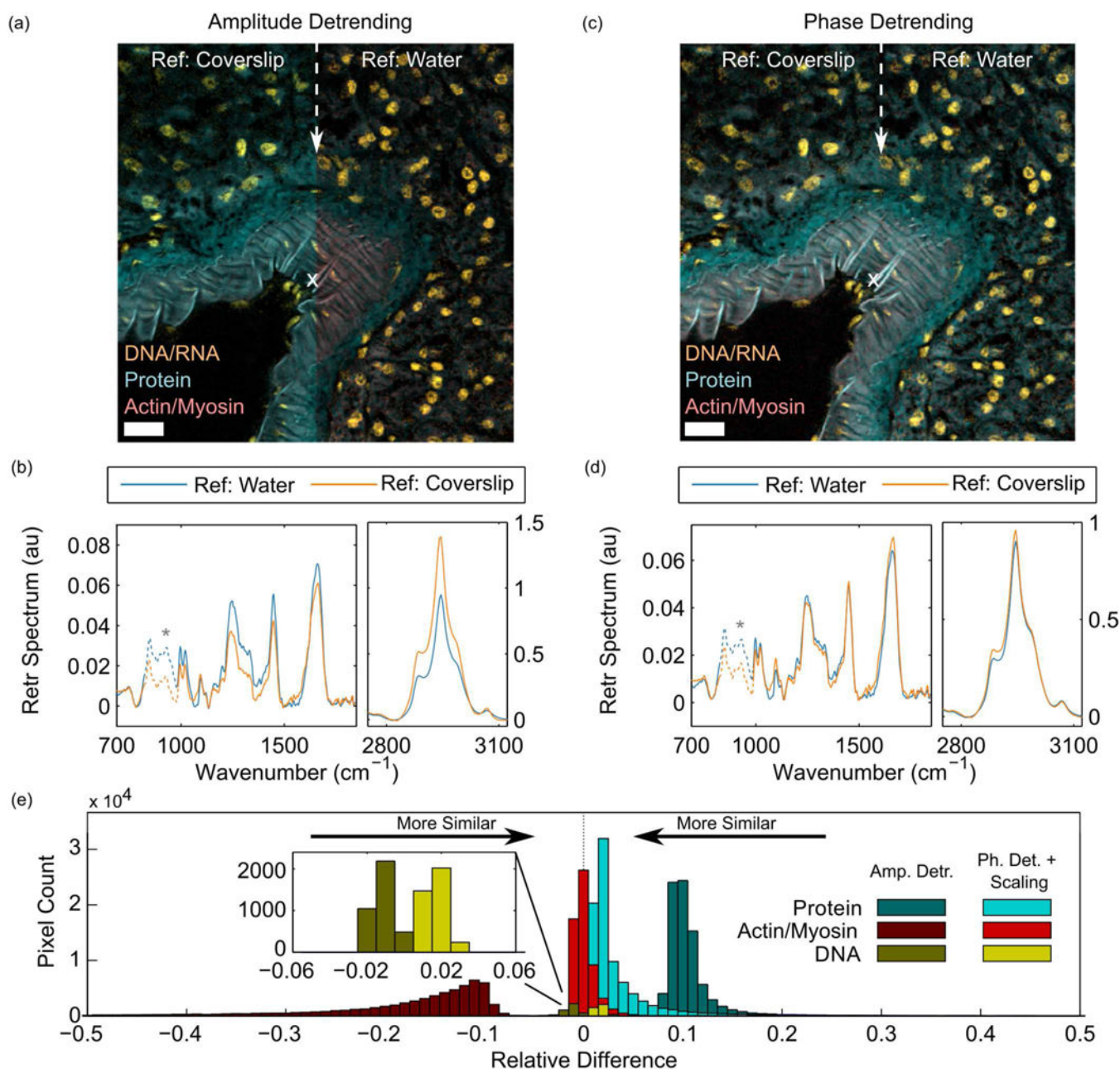


Figure 5.

Spectral retrieval and correction of hyperspectral data: murine pancreatic artery tissue section. (a) Traditional method of preparing pseudocolor imagery shows distinct differences when using a glass coverslip or water as the NRB reference, which cannot be simply corrected with normalization. (b) Single-pixel spectra (white 'x' in (a)) show intra-spectral deviations that are reference-dependent. Additional errors due to Raman peaks emanating from the reference NRB materials ('*' and dashed lines). (c) Pseudocolor imagery using the presented phase retrieval shows no obvious sign of difference between the halves processed with different NRB references. (d) Single-pixel spectra show close agreement with residual error due to Raman peaks emanating from the reference NRB materials ('*' and dashed

lines). (e) Histogram analysis of the relative difference between spectral peak intensities used in the creation of (a). The developed method shows <4% difference (mean) between peak amplitudes.

NIST Author Manuscript

NIST Author Manuscript

NIST Author Manuscript