

# Sparse sampling for fast hyperspectral coherent anti-Stokes Raman scattering imaging

Francesco Masia,<sup>1,\*</sup> Paola Borri,<sup>1,2</sup> and Wolfgang Langbein<sup>1</sup>

<sup>1</sup> Cardiff University School of Physics and Astronomy, The Parade, Cardiff CF24 3AA, UK

<sup>2</sup> Cardiff University School of Biosciences, Museum Avenue, Cardiff CF10 3AX, UK

[\\*masiaf@cardiff.ac.uk](mailto:*masiaf@cardiff.ac.uk)

**Abstract:** We demonstrate a method to increase the acquisition speed in coherent anti-Stokes Raman scattering (CARS) hyperspectral imaging while retaining the relevant spectral information. The method first determines the important spectral components of a sample from a hyperspectral image over a small number of spatial points but a large number of spectral points covering the accessible spectral range and sampling the instrument spectral resolution at the Nyquist limit. From these components we determine a small set of frequencies needed to retrieve the weights of the components with minimum error for a given measurement noise. Hyperspectral images with a large number of spatial points, for example covering a large spatial region, are then measured at this small set of frequencies, and a reconstruction algorithm is applied to generate the full spectral range and resolution. The resulting spectra are suited to retrieve from the CARS intensity the CARS susceptibility which is linear in the concentration, and apply unsupervised quantitative analysis methods such as FSC<sup>3</sup> [1]. We demonstrate the method on CARS hyperspectral images of human osteosarcoma U2OS cell, with a reduction in the acquisition time by a factor of 25. This method is suited also for other coherent vibrational microscopy techniques such as stimulated Raman scattering, and in general for hyperspectral imaging techniques with sequential spectral acquisition.

© 2014 Optical Society of America

**OCIS codes:** (180.4315) Nonlinear microscopy; (300.6230) Spectroscopy, coherent anti-Stokes Raman scattering.

---

## References and links

1. F. Masia, A. Glen, P. Stephens, P. Borri, and W. Langbein, "Quantitative chemical imaging using hyperspectral coherent anti-Stokes Raman scattering microscopy," *Anal. Chem.* (2013). DOI 10.1021/ac402303g.
2. C. L. Evans and X. S. Xie, "Coherent anti-Stokes Raman scattering microscopy: chemical imaging for biology and medicine," *Annu. Rev. Anal. Chem.* **1**, 883–909 (2008).
3. J. P. Pezacki, J. A. Blake, D. C. Danielson, D. C. Kennedy, R. K. Lyn, and R. Singaravelu, "Chemical contrast for imaging living systems: molecular vibrations drive CARS microscopy," *Nat. Chem. Biol.* **7**, 137–145 (2011).
4. A. Zumbusch, W. Langbein, and P. Borri, "Nonlinear vibrational microscopy applied to lipid biology," *Progress in Lipid Research* **52**, 615–632 (2013).
5. C. L. Evans, E. O. Potma, M. Puoris'haag, D. Côté, C. P. Lin, and X. S. Xie, "Chemical imaging of tissue in vivo with video-rate coherent anti-Stokes Raman scattering microscopy," *Proc. Natl. Acad. Sci. U. S. A.* **102**, 16807–16812 (2005).
6. C. W. Freudiger, W. Min, B. G. Saar, S. Lu, G. R. Holtom, C. He, J. C. Tsai, J. X. Kang, and X. S. Xie, "Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy," *Science* **322**, 1857–1861 (2008).

7. Y. Ozeki, F. Dake, S. Kajiyama, K. Fukui, and K. Itoh, "Analysis and experimental assessment of the sensitivity of stimulated Raman scattering microscopy," *Opt. Express* **17**, 3651–3658 (2009).
8. B. G. Saar, C. W. Freudiger, J. Reichman, C. M. Stanley, G. R. Holtom, and X. S. Xie, "Video-rate molecular imaging in vivo with stimulated Raman scattering," *Science* **330**, 1368–1370 (2010).
9. D. Fu, F.-K. Lu, X. Zhang, C. Freudiger, D. R. Pernik, G. Holtom, and X. S. Xie, "Quantitative chemical imaging with multiplex stimulated Raman scattering microscopy," *Journal of the American Chemical Society* **134**, 3623–3626 (2012).
10. I. Pope, W. Langbein, P. Watson, and P. Borri, "Simultaneous hyperspectral differential-CARS, TPF and SHG microscopy with a single 5 fs Ti:Sa laser," *Opt. Express* **21**, 7096–7106 (2013).
11. L. Kong, M. Ji, G. R. Holtom, D. Fu, C. W. Freudiger, and X. S. Xie, "Multicolor stimulated Raman scattering microscopy with a rapidly tunable optical parametric oscillator," *Opt. Lett.* **38**, 145–147 (2013).
12. Y. Liu, Y. J. Lee, and M. T. Cicerone, "Broadband CARS spectral phase retrieval using a time-domain Kramers-Kronig transform," *Opt. Lett.* **34**, 1363–1365 (2009).
13. E. M. Vartiainen, H. A. Rinia, M. Müller, and M. Bonn, "Direct extraction of Raman line-shapes from congested CARS spectra," *Opt. Express* **14**, 3622–3630 (2006).
14. D. Zhang, M. N. Slipchenko, D. E. Leaird, A. M. Weiner, and J.-X. Cheng, "Spectrally modulated stimulated Raman scattering imaging with an angle-to-wavelength pulse shaper," *Opt. Express* **21**, 13864–13874 (2013).
15. Y. Ozeki, W. Umemura, Y. Otsuka, S. Satoh, H. Hashimoto, K. Sumimura, N. Nishizawa, K. Fukui, and K. Itoh, "High-speed molecular spectral imaging of tissue with stimulated Raman scattering," *Nature Photon.* **6**, 845–851 (2012).
16. S. Qaisar, R.M. Bilal, W. Iqbal, M. Naureen, and S. Lee, "Compressive sensing: from theory to applications, a survey," *J. Communications and Network* **15**, 443–456 (2013).
17. T. Hellerer, A. M. Enejder, and A. Zumbusch, "Spectral focusing: High spectral resolution spectroscopy with broad-bandwidth laser pulses," *Appl. Phys. Lett.* **85**, 25–27 (2004).
18. I. Rocha-Mendoza, W. Langbein, and P. Borri, "Coherent anti-Stokes Raman microspectroscopy using spectral focusing with glass dispersion," *Appl. Phys. Lett.* **93**, 201103 (2008).
19. W. Langbein, I. Rocha-Mendoza, and P. Borri, "Coherent anti-Stokes Raman micro-spectroscopy using spectral focusing: Theory and Experiment," *J. Raman Spectrosc.* **40**, 800–808 (2009).

## 1. Introduction

In the last decade coherent Raman scattering (CRS) microscopy has emerged as important label-free and chemically sensitive imaging technology for biological systems [2–4]. Coherent Raman scattering probes intrinsic vibrational resonances to provide chemical contrast, which alleviates the need of perturbing the natural environment of biomolecules by the addition of labels such as fluorophores used for fluorescence imaging.

The two most common CRS imaging techniques are coherent anti-Stokes Raman scattering (CARS) [2, 3, 5] and stimulated Raman scattering (SRS) [6–9]. In both methods, the coherent vibration of the biomolecules is driven by the interference between two electromagnetic fields called pump and Stokes fields. The pump field is then Raman scattered by the driven vibrations. In CARS, the anti-Stokes part of the scattered field is acquired which is free from excitation background, while in SRS the scattering is measured by the loss of the pump beam or the gain of the Stokes beam. The nonlinear character of CARS and SRS enables sub-micron three dimensional resolution and the coherent superposition of the scattering of all molecules in the focal volume yields a significantly increased signal compared to spontaneous Raman scattering, allowing for fast imaging.

In both techniques, the measurement of a vibrational spectrum over a significant number of spectral points is required to provide a good chemical specificity in the presence of several chemical compounds. Such hyperspectral vibrational microscopy can be achieved by sequential imaging at different vibrational frequencies [10, 11]. For CARS intensity data, measuring the vibrational spectrum over a sufficiently large spectral range is actually required to extract quantitative chemical information. This is due to the fact that the quantity which is linear in the chemical composition is the complex susceptibility, while the CARS intensity is proportional to the absolute square of the susceptibility, having a modified spectrum compared to spontaneous Raman [4]. Several methods are available to retrieve the complex susceptibility spectra from

CARS intensity spectra [1, 12, 13]. These methods determine the phase of the susceptibility from the CARS intensity, which is a spectrally non-local operation, and consequently the resulting errors decrease with increasing CARS spectral range [1]. Such a retrieval method is not required for SRS which is proportional to the imaginary part of the susceptibility. However, the SRS signal can contain significant background due to Kerr-lensing, thermal lensing, or induced absorption [14]. Furthermore, we have recently shown that the real part of the susceptibility measured in CARS provides an internal calibration standard, and can also be used to distinguish chemical substances which do not show vibrational resonances in the measured spectral range [1].

To date, the fastest SRS imaging was shown in [15], acquiring images of  $500 \times 500$  spatial pixel at a rate of 30 frame per seconds (fps) with frame-by-frame frequency tunability. Considering the 91 frequencies acquired in the range  $2800\text{-}3100\text{ cm}^{-1}$  and 10 repetitions needed for sufficient signal-to-noise ratio, this corresponds to 0.03 hyperspectral frames per second (hfps). CARS imaging can reach similar speeds [2, 5]. Significantly faster rates are required for real-time imaging with quantitative chemical information, and for high-throughput high content screening, where a large number of similar structures need to be measured to provide statistically relevant results.

Here we demonstrate a method to reduce the number of measured frequencies in the hyperspectral image, and consequently to increase the hfps rate, without losing significant spectral information. Such sparse sampling methods reconstructing signals from far fewer samples than Nyquist sampling have recently attracted considerable attention [16], and applications cover many areas, such as geological remote sensing and medical imaging. Our work is developing and applying a sparse sampling method for coherent Raman hyperspectral imaging.

## 2. Experimental

CARS hyperspectral images have been acquired on a home-built multi-modal laser-scanning microscope based on an inverted Nikon Ti-U. A detailed description of the set-up can be found in [10]. A broadband (660-970 nm) laser beam from a 5 fs Ti:Sa laser is split into pump and Stokes beams for CARS excitation with wavelength ranges of 660-730 nm and 730-900 nm, respectively. Hyperspectral imaging is achieved by spectral focussing [17–19]. The pump and Stokes pulses are equally linearly chirped, resulting in a constant instantaneous frequency difference (IFD) within the pulse duration. Excitation of different vibration energies in the range  $1200\text{-}3800\text{ cm}^{-1}$  with a resolution of  $10\text{ cm}^{-1}$  is achieved by controlling the delay between pump and Stokes. The data discussed in this paper were taken over a  $(2400\text{-}3700)\text{ cm}^{-1}$  range with a  $60 \times 1.27$  NA water immersion objective (Nikon CFI Plan Apochromat IR  $\lambda$ S series) and a 1.4 NA oil immersion condenser for signal collection in forward direction, with a resulting spatial resolution for the CARS susceptibility (full-width at half-maximum of the point-spread function amplitude) of  $0.6 (2)\ \mu\text{m}$  in the lateral (axial) direction. The CARS light is discriminated by a pair of Semrock band-pass filters FF01-562/40 and then detected by a Hamamatsu H7422-40 photomultiplier. CARS hyperspectral images of a human osteosarcoma U2OS cell were acquired for IFDs in the range  $(2400\text{-}3700)\text{ cm}^{-1}$  with  $5\text{ cm}^{-1}$  steps. An example image at  $2805\text{ cm}^{-1}$  is shown in Fig. 1.

## 3. Results and discussion

The method proceeds as follows. First a hyperspectral data set  $A$  with a large spectral range and high spectral resolution is taken over a small number of spatial points. Singular value decomposition is used to determine the important spectra above the noise in  $A$ . Then a number of important spectral points equal to the number of important spectra is determined to yield minimum deviation over  $A$  in the spectra reconstructed from the important spectral points.

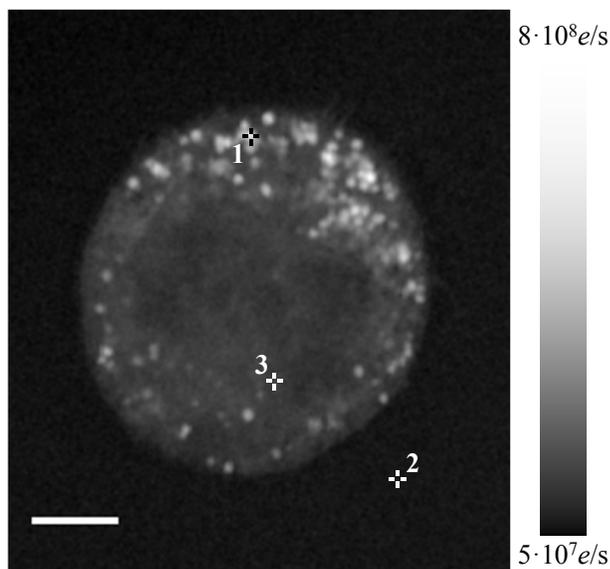


Fig. 1. CARS intensity  $I_C$  image of a human osteosarcoma U2OS cell acquired at an IFD of  $2805\text{ cm}^{-1}$ . Logarithmic gray scale as shown. The pixel size is  $0.1 \times 0.1\ \mu\text{m}^2$ , and the pixel dwell time was  $10\ \mu\text{s}$ . The pump and Stokes beam power at the sample were  $53\text{ mW}$  and  $31\text{ mW}$ , respectively. Spectra at the positions indicated in the Fig. are shown in Fig. 3. The scale bar represents  $5\ \mu\text{m}$ . The full dataset of  $\sqrt{I_C}$  is shown in the panel a) of the hyperspectral movie [Media 1](#).

Subsequently, hyperspectral data sets over large number of spatial points taken only at the important spectral points can be reconstructed.

### 3.1. Determination of important spectra and spectral points

A hyperspectral CARS intensity ( $I_C$ ) image  $A$  with a spectral step size at the Nyquist-limit of the spectral resolution is acquired using either a coarse spatial sampling or a small region of the sample containing a representative ensemble of the chemical components present. The subsequent processing is done on the square root of  $I_C$ , which has a whitened shot noise [1]. The  $\sqrt{I_C}$  data are formatted as a  $(S \times P')$  matrix with the elements  $\mathbf{d}_{s,p}$ , where  $S$  is the number of spectral points and  $P'$  is the number of spatial points measured. The matrix  $\mathbf{d}$  is then factorized using a compact singular value decomposition (SVD) into  $\mathbf{d} = \mathbf{u}\boldsymbol{\sigma}\mathbf{v}^T$ , where the  $(S \times S)$  matrix  $\mathbf{u}$  is the rotation from the new spectral basis given by the singular spectra, to the old spectral basis given by the measured IFDs. The  $(S \times S)$  matrix  $\boldsymbol{\sigma}$  is diagonal containing the singular values, and the  $(S \times P')$  matrix  $\mathbf{v}^T$  contains the spatial images in the basis of the singular vectors. As example we use as hyperspectral data  $A$  the data shown in Fig. 1 having  $S = 261$ , but with a coarse spatial step of  $1\ \mu\text{m}$ , containing only 1% of the spatial points of the frame, yielding  $P' = 951$ .

We determine the number  $S_{\text{max}}$  of singular values not dominated by noise using the algorithm reported in [1]. It consists of fitting the singular values  $\sigma_{j,j}$  (see Fig. 2(a) with a linear dependence on the index  $j$  for  $j \geq S/2$ , and determine the largest index  $j = S_{\text{max}}$  for which  $\sigma_{j,j}$  is larger than  $\sqrt{2}$  times the fit. For the example in Fig. 2(a) we find  $S_{\text{max}} = 8$ . Components with  $j > S_{\text{max}}$  are attributed to noise and we discard them by defining a  $\tilde{\boldsymbol{\sigma}}$  in which the diagonal

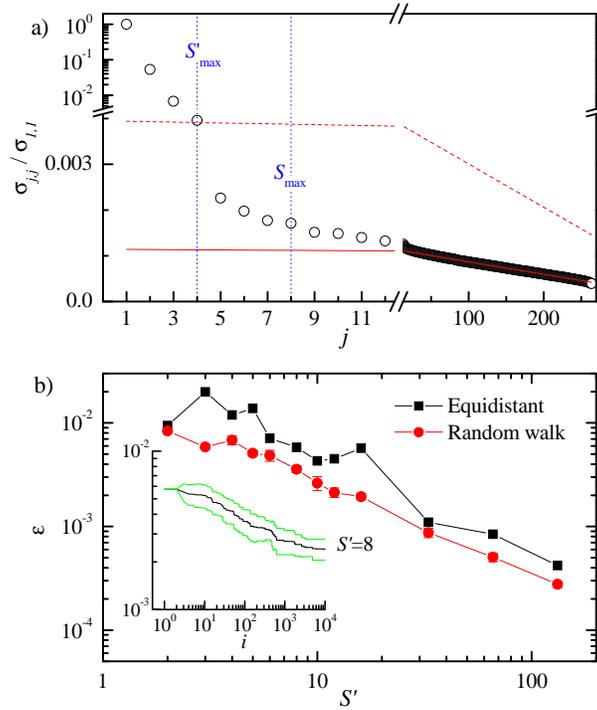


Fig. 2. a) Diagonal elements of the singular value matrix  $\sigma$  of dataset  $A$ , normalized to the largest value. The solid line is a linear fit to  $\sigma_{j,j}$  for  $j \geq S/2$ . The dashed line is the linear fit multiplied by  $\eta$ . b) Error  $\varepsilon$  in the spectral reconstruction of the data as a function of the number  $S'$  of spectral points used, for the different methods discussed in the text as labeled. The inset shows the evolution of  $\varepsilon$  during the random walk for  $S' = 8$ , where  $i$  indicates the step number. The green surrounding lines bars indicate the standard deviation of  $\varepsilon$  over 10 independent repetitions of the walk.

values are set to zero for  $j > S_{\max}$  and define the noise-filtered data  $\tilde{\mathbf{d}} = \mathbf{u}\tilde{\sigma}\mathbf{v}^T$ .

In order to reduce the number of spectral points to be measured from  $S$  to  $S' \ll S$ , we use only the most important  $S'_{\max} \leq S_{\max}$  singular spectra, and determine their weights in the reconstruction by measurements at  $S' \geq S'_{\max}$  spectral points.  $S'_{\max}$  and the specific spectral points to be measured are determined from  $\tilde{\mathbf{d}}$ . In detail, we create a  $(S' \times P')$  matrix  $\hat{\mathbf{d}}$  by selecting  $S'$  out of the  $S$  spectral points of  $\tilde{\mathbf{d}}$ , which are given by a vector  $\mathbf{s} = (s_1, \dots, s_{S'})$  of ascending spectral point indices, i.e.  $\hat{\mathbf{d}}_{n,p} = \tilde{\mathbf{d}}_{s_n,p}$  with  $n = 1..S'$ . An approximation of the matrix  $\tilde{\mathbf{d}}$  can be obtained by determining the weights of the  $S'_{\max}$  most important singular spectra using  $\hat{\mathbf{d}}$ . To do this, we determine a  $S' \times S'$  rotation matrix  $\hat{\mathbf{u}}$  by taking the first  $S'$  singular spectra of  $\mathbf{u}$  at the spectral points identified by  $\mathbf{s}$ , and using Gram-Schmidt orthogonalization of the resulting matrix in sequence of increasing singular spectrum index, i.e. decreasing importance. We then calculate a reduced matrix of coefficients  $\hat{\mathbf{a}}$  such as  $\hat{\mathbf{d}} = \hat{\mathbf{u}}\hat{\mathbf{a}}$ . Having the coefficients of the singular spectra, we can reconstruct the data over all spectral points using the  $(S \times S')$  sub-matrix  $\tilde{\mathbf{u}}$  of  $\mathbf{u}$  containing the first  $S'$  singular spectra. Similar to the SVD noise filtering, we restrict the reconstruction to  $S'_{\max} \leq S'$  important spectra by setting in  $\tilde{\mathbf{u}}$  the singular spectra above  $S'_{\max}$  to zero. The noise filtered reconstructed data is then given by  $\mathbf{d}^{\text{rec}} = \tilde{\mathbf{u}}\hat{\mathbf{u}}^{-1}\hat{\mathbf{d}}$

To determine  $S'_{\max}$  for the reconstruction, we have to consider that  $S_{\max}$  was determined using the noise in a dataset with  $S$  spectral components. Measuring only  $S'$  spectral points, the noise

is expected to be a factor of  $\sqrt{S/S'}$  larger. Furthermore, we note that we can choose the spectral points  $\mathbf{s}$  to minimize the influence of noise in the reconstruction. We accordingly determine  $\mathbf{s}$  by minimizing the error of the reconstructed hyperspectral image  $\varepsilon = \|\tilde{\mathbf{d}} - \mathbf{d}^{\text{rec}}\|_{\text{F}} / \|\tilde{\mathbf{d}}\|_{\text{F}}$ , where  $\|\cdot\|_{\text{F}}$  indicates the Frobenius norm, and  $S'_{\text{max}} = S_{\text{max}}$  is used to determine  $\mathbf{d}^{\text{rec}}$ . For a typical number of  $S$  in the order of 100, and  $S'$  in the order of 10, the number of possible combinations in  $\mathbf{s}$  is  $S!/(S'!(S-S')!) \sim 10^{13}$ , making it computationally unfeasible to find the exact minimum by calculating the error for each combination. We have therefore employed the following minimization algorithm. We start by choosing equidistant spectral points in  $\mathbf{s}$ , which corresponds to an under-sampling of the spectral resolution, and we choose the  $\mathbf{s}$  with minimum distance and centered in the spectral range. The minimization is then done by a random walk, choosing new spectral points  $\mathbf{s}'$  with  $s'_i = s_i + \delta_i$  using a random value  $\delta_i$  having a uniform distribution in the range  $s_{i-1} - s_i < 2\delta_i < s_{i+1} - s_i$ , i.e. covering half the distance to the adjacent spectral points, such that the ascending order is conserved. The new spectral points  $\mathbf{s}'$  are kept for subsequent steps only if the error improved. We limit the number of steps in the random walk, to 200 in the results shown here. We also investigated other minimization algorithms leading generally to similar or inferior results in terms of convergency and computational complexity. The error  $\varepsilon$  of the reconstruction for equidistant spectral points,  $\varepsilon^{\text{eq}}$ , and for random walk optimized points,  $\varepsilon^{\text{rw}}$ , is shown in Fig. 2(b). The error reduces with increasing number of measured spectral points  $S'$  approximately as  $1/\sqrt{S'}$ .

The evolution of  $\varepsilon$  during the random walk minimization is shown in the inset of Fig. 2(b) for  $S' = S_{\text{max}} = 8$ . Most of the improvement is realized within  $10^3$  steps. The standard deviation of  $\varepsilon$  over 10 independent walks is generally below  $0.2\varepsilon$ , so that we can conclude that the minimization algorithm is sufficiently stable. Interestingly, the error for equidistant spectral points is typically only a factor of two larger than after random walk minimization, showing that the exact choice of the spectral points is in general not important for the retrieval.

The number of important components  $S'_{\text{max}}$  for the reconstruction is then determined taking into account the modification of the noise by multiplying the linear fit of the noise components in  $\sigma$  (see solid line in Fig. 2) by the factor  $\eta = (\varepsilon^{\text{rw}}/\varepsilon^{\text{eq}})\sqrt{S/S'}$ , and is shown as dashed line in Fig. 2(a).

Spectral components which are not represented in the SVD basis, i.e. which are within the noise of the small spatial dataset, are not retrieved by the sparse sampling method.

### 3.2. Reconstruction of the spectra of a sparsely sampled hyperspectral image

Having determined  $\mathbf{s}$ , a hyperspectral CARS intensity dataset  $B$  is acquired at the  $S'$  spectral points of  $\mathbf{s}$  over a large spatial region with a number of spatial points  $P \gg P'$ . We used here the image shown in Fig. 1, for which  $P = 95120$ , and we chose  $S' = S_{\text{max}} = 8$ , for which we find  $S'_{\text{max}} = 4$  (see Fig. 2). We format the data in  $B$  as a  $S' \times P$  matrix  $\hat{\mathbf{D}}$  of  $\sqrt{I_C}$  and reconstruct the data over all spectral points using  $\mathbf{D}^{\text{rec}} = \tilde{\mathbf{u}}\hat{\mathbf{u}}^{-1}\hat{\mathbf{D}}$ .

To evaluate the retrieval, we compare in Fig. 3(a) the spectra of  $I_C$  at the positions marked in Fig. 1 as obtained from the reconstruction algorithm (dashed lines) and directly measured and denoised (solid lines). The reconstruction reproduces the measured spectra up to small deviations, which are more important for the position #1 corresponding to a lipid droplet in the cell. We have verified that using  $S_{\text{max}}$  instead of  $S'_{\text{max}}$  components in the reconstruction results in a significantly increased noise in  $\mathbf{D}^{\text{rec}}$ , as illustrated in Fig. 3(a) (gray line). The full datasets of the denoised and reconstructed  $\sqrt{I_C}$  are shown in panel b) and c) of the hyperspectral movie [Media 1](#), respectively.

To determine the CARS susceptibility, we have measured the CARS intensity  $I_{\text{ref}}$  in a non-resonant medium (the coverslip glass) to calculate the CARS ratio  $\tilde{I}_C = I_C/I_{\text{ref}}$ , which is then used to retrieve the imaginary part of the complex susceptibility  $\Im(\tilde{\chi})$  using the phase-corrected

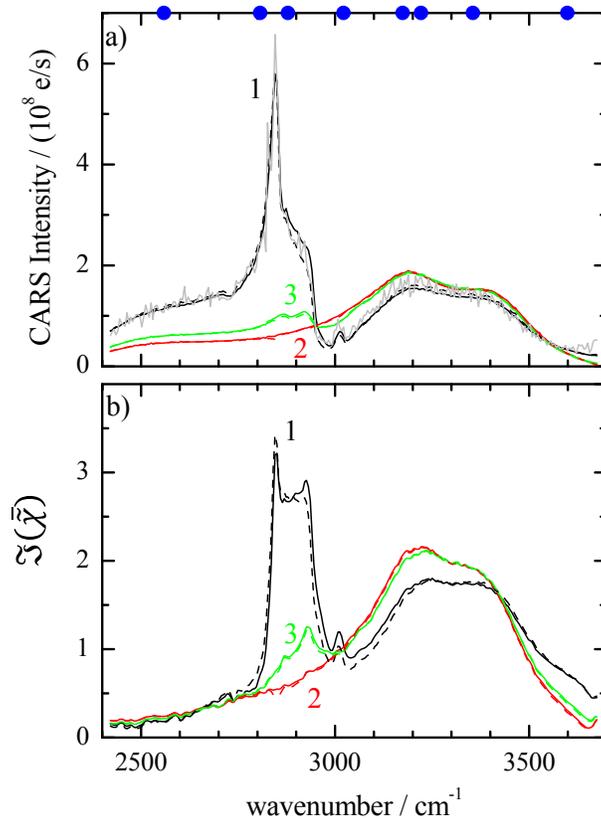


Fig. 3. (a) Comparison of the measured (solid lines) and reconstructed (dashed line) CARS intensity spectra at the positions indicated in Fig. 1. The light gray spectrum is obtained using  $S'_{\max} = S_{\max}$ . The blue dots represent the  $S' = 8$  wavenumbers given by  $\mathbf{s}$  used in the reconstruction of the hyperspectral image. A hyperspectral movie of the measured and reconstructed  $\sqrt{I_C}$  is shown in panels b) and c) of [Media 1](#), respectively. (b) Corresponding retrieved  $\Im(\tilde{\chi})$  using the PCKK method. A hyperspectral movie of the measured and retrieved  $\Im(\tilde{\chi})$  is shown in panels b) and c) of [Media 2](#), respectively.

Kramers-Kronig (PCKK) method described in [1]. In Fig. 3 we compare the resulting  $\Im(\tilde{\chi})$  at the positions marked in Fig. 1 as obtained from the reconstruction (dashed lines) and directly measured (solid lines). We see that the reconstructed  $\Im(\tilde{\chi})$  reproduces the directly measured spectra up to small deviations. The full datasets of the retrieved  $\Im(\tilde{\chi})$  using unfiltered, SVD denoised or reconstructed  $\tilde{I}_C$  are shown in panels a), b) and c) of the hyperspectral movie [Media 2](#), respectively.

We reconstruct  $\sqrt{I_C}$  because it has a white noise and spectrally local to the measured  $I_C$ . One could also reconstruct  $\log(\tilde{\chi})$ , since it is linearly related to  $\log(\tilde{I}_C)$  via PCKK, but it is spectrally non-local and not of white noise.

To investigate the error of the reconstruction in more detail, we consider the spatial distribution of the spectral error  $E_{S_p} = \sqrt{P} \left( \sum_{j=1}^S (\mathbf{D}_{j,p} - \mathbf{D}_{j,p}^{\text{rec}})^2 \right) / \|\mathbf{D}\|_F$  where  $\mathbf{D}$  is the noise filtered data taken at full spectral and spatial resolution. Figure 4 shows the spatial dependence of the spectral error for  $\sqrt{I_C}$  and  $\Im(\tilde{\chi})$ . The largest deviations can be observed at the lipid droplets.

The demonstrated sparse sampling method corresponds to an improvement in the acquisition

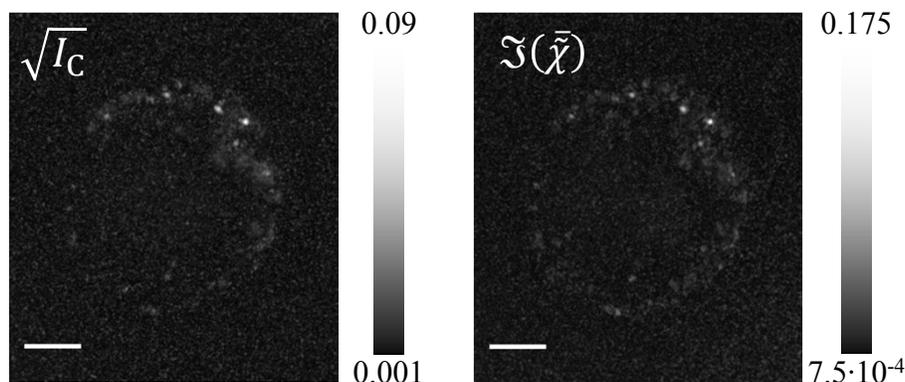


Fig. 4. Spectral error  $E_S$  of the  $\sqrt{I_C}$  and  $\mathfrak{S}(\tilde{\chi})$  due to reconstruction from sparse sample data a linear grayscale as shown. The scale bar represents  $5 \mu\text{m}$ .

speed by a factor  $(S_{\text{max}}/S + \rho)^{-1}$  where  $\rho = P'/P$  is the fraction of spatial positions in dataset  $A$  compared to  $B$ . For the example shown here we have  $S_{\text{max}} = 8$ ,  $S = 261$ ,  $\rho = 0.01$  resulting in a 25 fold increase in acquisition speed. In our analysis it is required that  $\rho P \gg S$  for a correct SVD-based noise filtering procedure. In typical situations where samples of a given chemical variety are imaged over a large set of replica, such as in high throughput screening, dataset  $A$  needs to be acquired only once, such that the speedup is effectively only limited by  $S/S'$ , which is 33 in the present case. In absolute terms it is typically sufficient to measure at 5-10 spectral points to reconstruct most of the chemical information obtainable in a full spectral scan.

#### 4. Conclusion

We have developed and demonstrated a sparse sampling method to reduce the acquisition time in CARS hyperspectral imaging. The method employs an adapted factorization of a large hyperspectral data matrix ( $S \times P$ ) into two smaller matrices of ( $S \times P'$ ) and ( $S' \times P$ ), with  $S' \ll S$  and  $P' \ll P$ . We demonstrated that the acquisition time for a human osteosarcoma U2OS cell was reduced by a factor of 25 without significant loss of information. This method, combined with state of the art CARS microscopes, is expected to enable real-time chemical imaging and high-throughput high-content label-free microscopy. The method is suited also for other coherent vibrational microscopy techniques such as stimulated Raman scattering, and in general for hyperspectral imaging techniques with sequential spectral acquisition.

#### Acknowledgments

This work was funded by the UK EPSRC Research Council (grant n. EP/H45848/1). The CARS microscope development was funded by the UK BBSRC Research Council (grant n. BB/H006575/1). FM acknowledges financial support from the European Union (Marie Curie Grant Agreement No. PERG08- GA-2010-276807). PB acknowledges the UK EPSRC Research Council for her Leadership fellowship award (grant n. EP/I005072/1). We thank Rachel Errington and Sally Chappell (Cardiff University) for providing the U2OS cells, Arnica Karuna for helping in the image acquisition, and Iestyn Pope for support in running the microscope. We acknowledge the European Union COST action MP1102 "Chemical imaging by Coherent Raman microscopy - microCoR" for supporting scientific exchange and discussions.