APL Photonics

RESEARCH ARTICLE | AUGUST 21 2024

# Thousand foci coherent anti-Stokes Raman scattering microscopy

Dominykas Gudavičius ; Lukas Kontenis ; Wolfgang Langbein



*APL Photonics* 9, 086110 (2024)
https://doi.org/10.1063/5.0220474

# Thousand foci coherent anti-Stokes Raman scattering microscopy

Cite as: APL Photon. 9, 086110 (2024); doi: 10.1063/5.0220474
Submitted: 25 May 2024 • Accepted: 30 July 2024 •
Published Online: 21 August 2024



Dominykas Gudavičius,[1,2] Lukas Kontenis,[1] and Wolfgang Langbein[2,a)]

## AFFILIATIONS

[1] Light Conversion, Keramiku st. 2B, LT-10233 Vilnius, Lithuania
[2] School of Physics and Astronomy, Cardiff University, The Parade, Cardiff CF24 3AA, United Kingdom

[a)] **Author to whom correspondence should be addressed:** langbeinww@cardiff.ac.uk

## ABSTRACT

We demonstrate coherent anti-Stokes Raman scattering (CARS) microscopy with 1089 foci, enabled by a high repetition rate amplified oscillator and an optical parametric amplifier. We employ a camera as a multichannel detector to acquire and separate the signals from the foci, rather than using the camera image itself. This allows us to retain the insensitivity of the imaging to scattering afforded by the non-linear excitation point-spread function, which is the hallmark of point-scanning techniques. We show frame rates of 0.3 Hz for a megapixel CARS image, limited by the camera used. The laser source and corresponding CARS signal allows for at least 1000 times higher speed, and using faster cameras would allow acquiring at that speed, opening a perspective to megapixel CARS imaging with a frame rate of more than 100 Hz.





## I. INTRODUCTION

Coherent Raman scattering (CRS) allows microscopy with chemical selectivity using the vibrational response of the chemical components of the sample without adding fluorescent dyes.[1–4] Driving the vibrations by an optical excitation which is intensity modulated at the vibrational frequency, the vibrations are, in turn, modulating the polarizability, creating coherent sidebands of the optical excitation fields. These sidebands are the CRS, and their amplitude and phase encode the vibrational spectrum of the medium probed. Since its revival about 2 decades ago,[5] enabled by the development of suited commercial laser sources, this field has developed a wide range of methods of excitation and detection, with different advantages and drawbacks.[1,4]

The most common scheme uses two synchronized optical pulses, the pump pulse and the longer wavelength Stokes pulse, with a frequency difference given by the vibrational frequency of interest. To achieve high intensities while keeping the average power below sample damage, pulses of picosecond duration, comparable to the typical vibrational coherence times in condensed matter, with repetition rates in the 1–100 MHz range are employed, having a duty cycle in the $10^{-4}$ to $10^{-6}$ range. Focusing these pulses to the diffraction limit of high numerical aperture (NA) objectives of about 0.5 $\mu$m size enables achieving peak intensities of the order of $10^{11}$–$10^{12}$ W/cm$^2$ for average powers of tens of milliwatts, suitable for efficient CRS generation just below typical saturation levels[6] around $10^{13}$ W/cm$^2$. Scanning this focus across the sample is then used to provide an image, similar to other multiphoton techniques recording two-photon fluorescence (TPF), second harmonic generation (SHG), or third harmonic generation (THG). The non-linearity of the CRS signal with the excitation intensity enables optical sectioning without imaging in the detection, an advantage compared to confocal or other structured illumination microscopes, particularly for inhomogeneous samples exhibiting significant diffuse scattering.

The high frequency sideband of the pump, called coherent anti-Stokes Raman scattering (CARS), and the low frequency sideband of the Stokes, call coherent Stokes Raman scattering (CSRS), can be detected free of an excitation background, while the high frequency sideband of the Stokes interferes with the pump, and creates stimulated Raman loss (SRL), and the low frequency sideband of the pump interferes with the Stokes, and creates stimulated Raman gain (SRG). Notably, sidebands are also created by purely electronic non-linearities, such as the Kerr effect, as well as by vibrational

resonances at higher frequencies, providing a background to the resonant response. These non-resonant responses are in quadrature to the pump or Stokes beams in SRL and SRG, respectively, yielding in the lowest order only a change of phase without changing the amplitude. Thus, by detecting the transmitted power, which is insensitive to the phase, only resonant responses are measured, rendering SRL and SRG spectra similar to Raman spectra. However, the relative changes are typically very small, in the $10^{-4}$ to $10^{-7}$ range, so that high-frequency modulation schemes and shot-noise limited laser sources are required.

Detecting the CARS or CSRS power, free of excitation background, does not have these requirements, but the signal contains the interference between resonant and non-resonant responses, creating asymmetric resonance line shapes, complicating the interpretation of the contrast. To retrieve a signal linear in the concentration of the chemical components, the complex susceptibility can be determined by phase retrieval using the causality of the response and the measured dependence on the vibrational frequency.[7–9]

The point scanning approach has been optimized and is now limited by the maximum signal available due to signal saturation by the excitation of higher vibrational levels[6] or heating and allows for video rate imaging.[10–12] However, small absorbing regions can lead to sample damage at the high intensities and repetition rates used, both by heating and avalanche breakdown. Recently, there was some effort to use squeezed light[13] to reduce the shot-noise in SRL and SRG, with modest success (less than a factor of two in signal to noise) limited by the un-squeezing of the detected light by losses due to scattering in the sample and finite transmission of optical elements. Thus, the only clear way forward to higher speed is to distribute the excitation into many focal spots or to use a wide-field illumination. This requires suited laser sources providing the large pulse energies creating an intensity that is sufficiently high for efficient CRS generation over the extended excitation region. To limit the average power on the sample, the repetition rate of the laser pulses needs to be reduced, which also allows the laser-induced heating and electronic excitations to dissipate between pulses, increasing damage thresholds.

Early wide-field approaches used nanosecond pulsed lasers with repetition rates in the 1–100 Hz range and a non-colinear geometry for phase matching, enabling to choose between dark-field (phase-mismatched) and bright-field (phase-matched) imaging.[14] The method was applied to image 2D cell cultures,[15] but the coherent wide-field nature of the signal created significant interference artefacts, particularly relevant for thicker and more scattering samples, and furthermore did not provide sectioning. This drawback was addressed using random speckle illumination, which after speckle averaging resembles a spatially incoherent illumination and thus removes the interference artefacts.[16] Multimode optical fibers were used to generate the speckles, with the Stokes beam of a coherence length much larger than the phase delay distribution across the fiber modes retaining the spatial speckle pattern, while the pump beam of a shorter coherence length creating a spatio-temporal speckle pattern, which provides some speckle averaging already over a single pulse.[16] Furthermore, using a varying speckle pattern across many camera frames and evaluating the fluctuations, the sectioning capability was recovered, but with a random noise in the image scaling with the inverse root of the number of independent frames, a number limited by signal strength and camera technology. Notably, the image formation relies exclusively on the imaging of the signal onto the camera and is thus affected by scattering in the collection path. The absence of a suited technical solution measuring the weak SRL or SRG modulation with an imaging detector presently limits these wide-field techniques to CARS and CSRS.

Recently, the wide-field speckle CARS imaging concept was revisited,[17] using an optical parametric amplifier providing a much higher repetition rate (515 kHz). The Stokes speckle pattern was changed only between camera frames, while the pump speckle pattern was changed many times over a single camera frame to provide averaging. A reconstruction algorithm (Wiener filtering) was used to recover the spatial resolution of the retrieved speckle fluctuation image in-plane. Notably, the issue of the speckle noise requiring averaging over many frames and the reliance on imaging in the collection path remained.

Instead of wide-field illumination, increasing the number of foci for fast scanning is a known concept in confocal microscopy[18] (Nipkow disk scanners) and also in multiphoton microscopy[19,20] where camera imaging in the detection path was used to provide the lateral resolution. Sectioning is created by apertures in the detection imaging for confocal microscopy and by the focused excitation itself in multiphoton microscopy. The inherent trade-off between the number of foci used and the sectioning capability has been investigated.[21] The only multifocal CARS imaging reported[22] up top now employed two synchronized Ti:sapphire oscillators, delivering 5 ps pulses at 80 MHz repetition rate. A rotating microlens array created an average of seven focal spots moving across the sample at any given time, and CARS was imaged onto a camera, which, during the exposure time, averaged the signal over the microlens positions fully covering the imaged area. Due to the low pulse energy available, the number of foci and the speed was rather limited, and while an application to cell imaging was shown,[23] the method was not pursued further.

In the present work, we revisit the multi-focal CARS microscopy concept and demonstrate imaging with a much larger number of foci at high speed. The commercial availability of high pulse energy and high repetition rate laser systems, such as those employed in Ref. 17, enables the excitation of a large number of foci without compromising on the CARS efficiency. Importantly, we introduce the use of a camera as a multi-segment detector to suppress the detrimental influence of scattering in the detection imaging. Particularly, the availability of fast and low noise CMOS cameras allows the read out of the CARS image for every position of the foci. This enables the use of camera images to determine the signal created by each individual focus as a sum over a defined region for each focus, emulating the non-imaging single channel detectors employed in single point scanning techniques. The spatial resolution of the resulting CARS image is created by the multiphoton excitation process, as in single point scanning techniques, allowing for deeper penetration into scattering samples.

## II. METHODS

### A. Samples

Sample O providing a homogeneous reference was prepared by pipetting 10 $\mu$L of olive oil (Bio Basso by Basso Fedele & Figli srl) onto a glass coverslip (No. 1.5, 24 × 50 mm$^2$, Thermo Scientific) and was spread by adding a second equal coverslip for imaging.

Sample A providing bead arrays in water was prepared by pipetting 10 $\mu$L of 1% solid suspension polystyrene (PS) beads of 2 $\mu$m diameter (124, Phosphorex) onto a glass coverslip (No. 1.5, 24 × 50 mm$^2$, Thermo Scientific). The mixture was left to dry under ambient conditions, and 10 $\mu$L of distilled water was pipetted onto the coverslip and spread by adding a second equal coverslip for imaging.

Sample B providing a bead mixture in oil was prepared by mixing 10 $\mu$L of 1% solid suspension of polymethyl methacrylate (PMMA) beads of 20 $\mu$m diameter (MMA20K, Phosphorex) and 10 $\mu$L of 1% solid suspension PS beads of 10 $\mu$m diameter (118, Phosphorex) on a glass coverslip as above. The mixture was left to dry under ambient conditions for 30 mins. Subsequently, 10 $\mu$L of olive oil was pipetted onto the coverslip and spread across the sample by adding a second equal coverslip for imaging.

Sample C providing small beads to measure the point spread function (PSF) was prepared by pipetting 10 $\mu$L of 1% solid suspension PS beads of 0.5 $\mu$m diameter (108, Phosphorex) onto a coverslip as mentioned above and drying under ambient conditions.

Human skin sarcoma samples to demonstrate the imaging performance were prepared by fixing them with 10% formalin and dehydrating it with ethanol. The embedding was performed by using paraffin, and the sample was sectioned with a microtome. The sample was stained by using a standard eosin and hematoxylin staining procedure.[24]

### B. Experimental Setup

The experimental setup used is sketched in Fig. 1(a). An amplified laser system (PHAROS, Light Conversion) provides transform-limited 290 fs pulses of 1020 nm center wavelength at 200 kHz repetition rate and 20 W average power. 5 W of the beam are used as Stokes beam for CARS, while the remaining power is used for an optical parametric amplifier (OPA) (ORPHEUS-F, Light Conversion), which provides 130 fs pulses tunable from 650 to 900 nm with the average power of 1 W, used as a pump beam for CARS. The powers of pump and Stokes are controlled by attenuators consisting of a half-wave plate (Castech) and a Glan–Taylor polarizer (GT10-B, Thorlabs). The Stokes beam is transmitted through 60 mm of SF-66 glass blocks (II-VI incorporated) and an optical delay line (VT-80, PI) with silver retroreflectors (Eksma optics) to control the temporal overlap of pump and Stokes at the sample. Both beams are combined into the same spatial mode by a dichroic mirror (high reflection from 600 to 950 nm and high transmission from 970 to 1300 nm, Eksma Optics) and then passed through 145 mm of SF-57 glass blocks (II-VI incorporated) for spectral focusing.[25,26] The pump at 785 nm accrues 33350 fs$^2$ of group delay dispersion (GDD) stretching it to 723 fs, while the fixed Stokes accrues 34450 fs$^2$ GDD stretching it to 439 fs. The difference in GDD can be compensated using a commercially available pulse compressor (Light Conversion), which can add up to 1300 fs$^2$ GDD to the pump beam. Both beams then pass through a long-pass filter (FEL0700, Thorlabs) to remove components below 700 nm wavelength, spectrally overlapping with the CARS signal (mostly residuals of the white-light generation seed of the OPA). A beam expander (plano-concave lens of focal length $f$ = −50 mm and plano-convex lens of $f$ = 150 mm, Casix) increases the beam size from 3 to 9 mm. The beams are then focused by a silica microlens square array (15–811, Edmund Optics) of 10 × 10 mm$^2$ size, 300 $\mu$m pitch, and $f$ = 8.7 mm, creating 33 × 33 focal spots. The generated beamlets are collimated by an achromatic doublet (PO-PK-L3, II-VI incorporated) of $f$ = 100 mm and are merging at its back-focal plane, which is located at the x-y galvo-mirrors (6215H, Cambridge Technology) scanning the beam directions. The beamlets are then relayed with four-times magnification (closely spaced pair of AC300-100-B and a AC508-200-B-ML, Thorlabs), expanding the beamlets to 13.8 × 13.8 mm$^2$ square size and imaging the galvo-mirrors into the 15 mm diameter back focal plane of the excitation objective (Plan Apo Lambda 20 × 0.75 NA, MRD00205, Nikon), which focuses the beamlets into the sample.

The transmitted light is collected by a second objective (Plan Apo VC 20 × 0.75 NA, 1501–9398, Nikon) and then passes through a short-pass and a band-pass optical filter (FESH0750 and FB650-40, Thorlabs) to isolate the CARS signal. A camera lens of $f$ = 50 mm and $f$/2.8 aperture (NMV-50M23, Navitar) images the sample with a



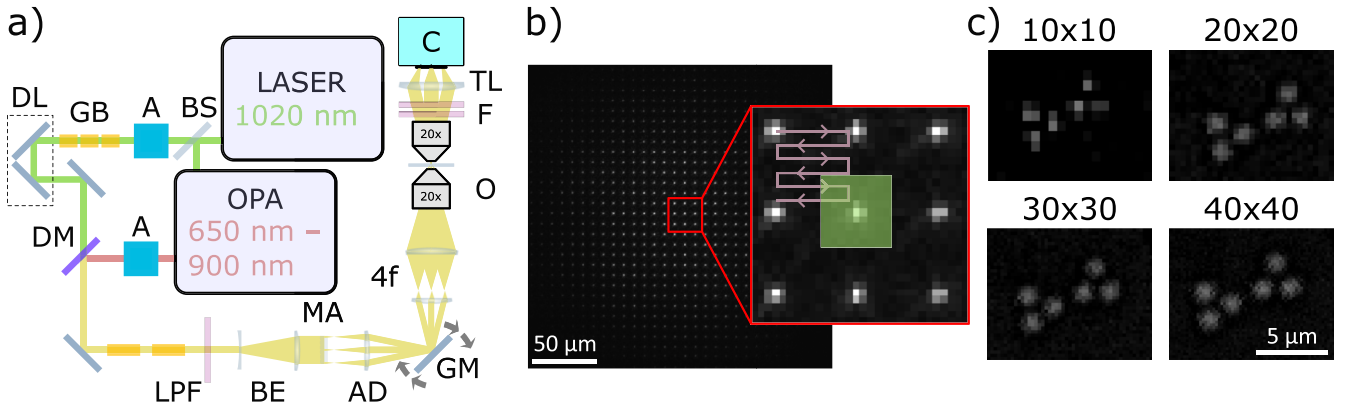


**FIG. 1.** (a) Multifocal CARS experimental setup. BS: beamsplitter; A: power attenuator; GB: glass blocks; DL: delay line; DM: dichroic mirror; LPF: long-pass filter; BE: beam expander; MA: microlens array; AD: achromatic doublet; GM: galvo mirrors; 4f: 4f relay system; O: objective; F: spectral filters; TL: tube lens; C: camera. (b) Multifoci raster-scan pattern. At each galvo-mirror position, an image of the signals from all the focal points is acquired. The pixel values within a square area (green) of $S_{int}$ = 10 pixels side length centered at each focal point are summed to provide the focal point signal. $g_c$ = 20 dB, $\tau_e$ = 4 ms, $\nu_f$ = 195 Hz, and the gray scale is black to white 0 to 1382 pe. (c) Examples of reconstructed images of 2 $\mu$m PS beads (sample A) for position grids of $N_g \times N_g$ points and $N_g$ of 10, 20, 30, and 40. The gray scale is black to white 0 to 1383 pe.

magnification of 5 onto a camera with 2448 × 2048 pixels of 3.45 $\mu$m size, $\sigma_r$ = 2.44 e read noise, 10 ke full well capacity, and USB3.1 interface (BFS–U3–51S5M-C, FLIR). The 18 mm aperture of the camera objective accommodates the 15 mm diameter of the collection objective back focal plane and allows for some off-axis beam walk-off. We note that the 5 times magnification onto the camera is slightly less than the 7 times required for Nyquist sampling, which reduces the number of pixels to be read per focal spot, improving the read noise and speed. The camera is triggered via the software driving the galvos with a frame rate $\nu_f$, and the exposure time $\tau_e$ is set to allow this frame rate, which requires $\tau_e + \tau_d < 1/\nu_f$ with the delay time $\tau_d$ = 75 $\mu$s. The camera is read with 8 bit pixel format for the highest speed, via a software written in LabView, and the frames are saved as bitmaps and processed off-line. We note that real-time processing of the data would be feasible with more software development. We used a camera analog gain $g_c$ of 10, 20, or 30 dB, providing 14.6, 4.61, or 1.46 e/count, respectively, and a digital offset of 5 counts to avoid zero-clipping. The dark signal was subtracted from each frame before further processing. Powers of excitation beams given later are measured at the sample.

We note that the laser repetition rate used is some 100 to 500 times lower than typically used in single point scanning techniques [1,2,3]. Thus, the pulse energy per focus for 1000 foci can be kept comparable to those techniques when using 2 to 10 times more power. Importantly, the pulse energy per focus dictates the CARS signal generation efficiency, and using established values close to saturation [6] for the optimal pulse durations in the 0.3–2 ps range, dictated by the vibrational coherence times, provides highest efficiencies. These can be reached with the present setup.

### C. Image acquisition

The CARS signal generated in the sample is imaged onto the camera, showing the 33 × 33 grid of beamlet foci, as exemplified in Fig. 1(b) for the homogeneous sample O. The intensities of the foci are varying across the array due to the variation of the beam intensity across the lenslet array considering the pump and Stokes excitation beams of Gaussian beam diameter (at $1/e^2$ intensity) of 9 mm, resulting in a gradual reduction toward the periphery of the array. Notably, CARS is a third-order signal proportional to $I_p^2 I_s$, with the pump intensity $I_p$ and the Stokes intensity $I_s$, resulting in a diameter of the CARS intensity of about 5 mm, a factor of $\sqrt{3}$ smaller than the beams. A refractive beam shaper could be used to provide a uniform intensity.

The galvo-mirrors are used to shift the foci in a raster-scan pattern with $N_g$ steps in each direction, filling the space between foci, as shown in Fig. 1(b). Their step response time of about 200 $\mu$s is well below the exposure times $\tau_e$. We note that this shift could also be achieved by simply laterally moving the microlens array. An image $P_{kl}$ is acquired at each raster-scan index $(k,l)$ in $x$ and $y$ directions for $k,l = 0,1,\ldots,N_g-1$. The foci are separated by 7.5 $\mu$m on the sample, which corresponds to $S_g$ = 10.9 pixels on the camera. The focal point positions $\mathbf{r}_{nm}$ of the beamlets of index $(n,m)$ in $x$ and $y$ directions, $n,m = 0,1,\ldots,32$, are determined using the signal peak positions in the image $P_{00}$ from the homogeneous sample. The camera centering and rotation is aligned so that the central row and column of foci are straight and deviating by less than half a pixel from the horizontal and vertical directions. A second-order polynomial $x(n) = x_0 + x_1 n + x_2 n^2$ is then fitted to the horizontal positions of the central row of peaks ($m$ = 16) and equivalently $y(m) = y_0 + y_1 m + y_2 m^2$ to the vertical positions of the central column of peaks ($n$ = 16), and the position of any peak is then taken as $\tilde{\mathbf{r}}(n,m) = (x(n), y(m))$ for image reconstruction. The second-order terms cater for small cushion distortions of the imaging. The nominal focus positions in $P_{kl}$ are taken as $\tilde{\mathbf{r}}(n + k/N_g, m + l/N_g)$ and are providing the signal for the pixel position $(i,j) = (nN_g + k, mN_g + l)$ in the excitation image, having a pixel size on the sample of 7.5 $\mu$m/$N_g$.

The pixel value $p_{ij}$ of the excitation image is evaluated by summing the pixel values of $P_{kl}$ over a square area with edge size $S_{int}$ centered on the nominal focus positions $\tilde{\mathbf{r}}$ [see the green area in Fig. 1(b)], noting that avoiding an overlap between foci requires $S_{int} < S_g$. The number of pixels summed is $S_{int}^2$, providing a read noise of $\sigma_s = S_{int}\sigma_r$, proportional to $S_{int}$. The variation of $p_{ij}$ with $S_{int}$ depends on the extension of the imaged signal in $P_{kl}$. In homogeneous samples, the extension is less than 2 pixels, as shown in Fig. 1(b), thus choosing $S_{int}$ = 3 seems appropriate to achieve the highest signal to noise. However, the extension is affected by sample inhomogeneities in the path from the beamlet focus to the camera, as exemplified in the supplementary material, Fig. S3. Thus, the immunity of $p_{ij}$ against sample inhomogeneities increases with $S_{int}$ and is maximized for $S_{int}$ = 10, as shown in Fig. S1. Excitation images of 2 $\mu$m PS beads (sample A) at 2950 cm$^{-1}$ are shown in Fig. 1(c) using different $N_g$, showing the increasing definition of the imaged beads with $N_g$. Nyquist sampling is achieved for $N_g \approx 30$, providing a pixel size of 250 nm.

The use of a camera to define a multi-segment detector separating the signal of each focus in the excitation imaging can be compared to the detection imaging employed in previous multifocal approaches,[19,22] where the image $P$ is given by the sum of $P_{kl}$ over all $(k,l)$. The resulting detection imaging is shown in Fig. 2(a) for sample A in a region with stacked patches of closely packed 2 $\mu$m PS beads in water, imaged at different axial planes. A significant blurring and reduced contrast in detection imaging (right column) compared to excitation imaging (left column) is observed. We note that the detection imaging pixel size of 0.69 $\mu$m is sufficient to resolve the beads. Thus, while the excitation imaging results in a PSF given by the excitation focus as in single point scanning, detection imaging results in blurring due to sample inhomogeneities, an issue that is intrinsic to the previous multifocal approaches.

### D. Image post-processing

Variations between microlenses lead to variations in the CARS signal from different foci. Moreover, these differences are modulated by sample inhomogeneities. Static, sample-independent variations can be corrected using an image of a homogeneous sample and calculating the ratio $r_{nm}$ of the average pixel value over the tile from an individual focus $A_{nm} = \langle p_{(nN_g+k,mN_g+l)} \rangle$ to the average value of all the tiles $\langle A_{nm} \rangle$. To exclude tiles of low signal, exhibiting excessive noise after correction, tiles with a ratio $r_{nm}$ lower than a threshold value $v$ are set to the average value of the other tiles $\bar{A}_v = \langle A_{nm} \rangle_{r_{nm}>v}$. The flat-field corrected pixel values are accordingly calculated as

$$\tilde{p}_{ij} = \begin{cases} p_{ij}/r_{nm}, & \text{if } r_{nm} \geq v, \\ \bar{A}_v, & \text{if } r_{nm} < v. \end{cases} \quad (1)$$

The uncorrected pixel values for the homogeneous sample O are shown in Fig. 2(b). A tile structure from imperfections in the

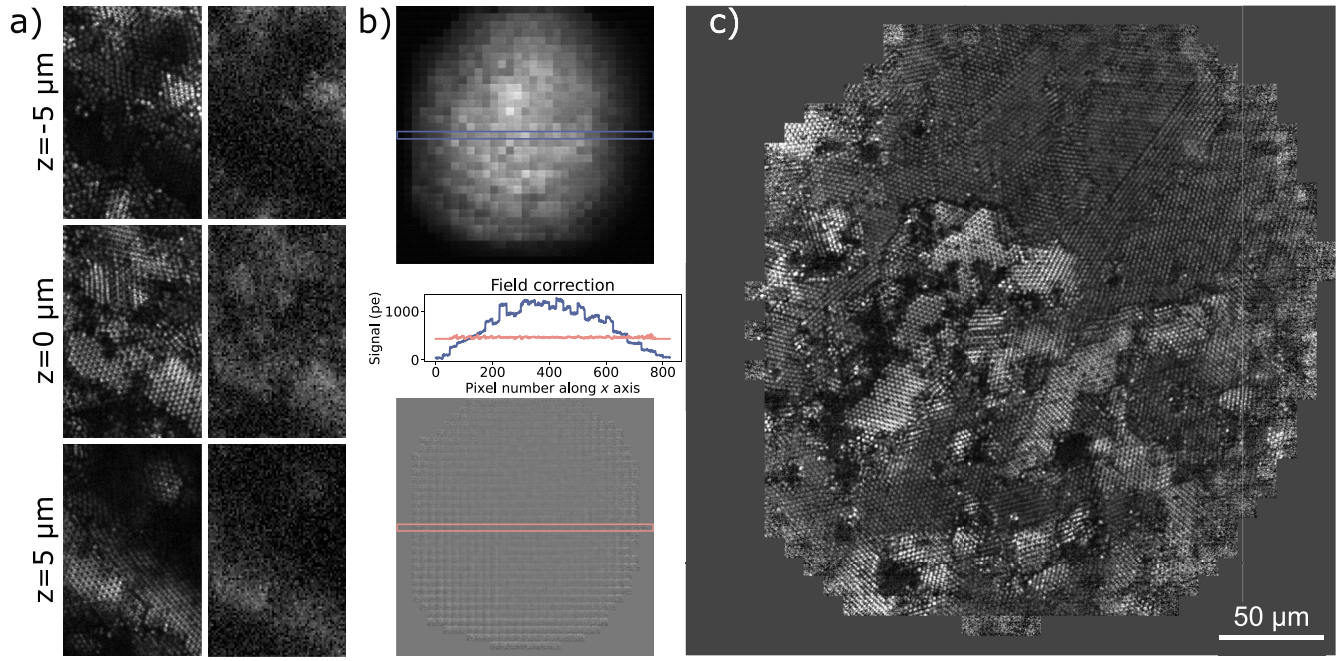


**FIG. 2.** (a) Excitation imaging (left) compared to detection imaging (right) at different $z$ planes as indicated. CARS of sample A (2 $\mu$m PS beads) at 2950 cm$^{-1}$ with 20 mW pump power and 50 mW Stokes power. $N_g = 30$, $\nu_f = 195$ Hz, $\tau_e = 4$ ms, $g_c = 30$ dB, $S_{int} = 8$, and greyscale is 0 to 1399 pe (left) and 0 to 19 575 pe (right). (b) Correction of beamlet-dependent signal strength. CARS imaging of the homogeneous sample O (oil) at 2850 cm$^{-1}$ with 20 mW pump power and 50 mW Stokes power. Top: image $p_{ij}$ shows a tile structure from imperfections in the microlens array, and vignetting by the excitation beam. $N_g = 25$, greyscale 0 to 1856 pe. The center cross section (blue line) is shown in the graph below. Applying the flat-field correction [Eq. (1)], the resulting image $\tilde{p}_{ij}$ is rather uniform (bottom, greyscale 0 to 1114 pe) as confirmed by the cross section (red line). (c) An example of a flat-field corrected image $\tilde{p}_{ij}$ of sample A, greyscale 0 to 466 pe, $\upsilon = 1/3$, settings as in panel (a).

microlens array and vignetting by the finite excitation beam size are visible. Applying the flat-field correction in Eq. (1), the resulting image $\tilde{p}_{ij}$ is rather uniform [see Fig. 2(b) at the bottom and red cross section]. There is a remaining gradient of some ±5% across the tiles, which is attributed to clipping of the excitation beams and the CARS signal by the objectives and could be reduced by optimizing the optical layout.

A corrected image of sample A with 2 $\mu$m PS beads is presented in Fig. 2(c), showing the high imaging quality achieved. Notably, even after correction, a tiling pattern is still visible. We attribute this to a distortion of the individual foci by the sample, which can reduce or increase the CARS signal depending on the distortions present without the sample due to the imperfections of the microlens array. Therefore, the relative intensity of the foci does depend on the refractive index inhomogeneities of the sample itself and cannot be fully compensated by a reference measurement on a homogeneous sample. These effects could be mitigated using a higher quality microlens array.

As an alternative to hardware improvements, we developed an algorithm to minimize the tile pattern *a posteriori*. This algorithm determines correction factors $C_{nm}$ for each beamlet to minimize the step heights at the tile edges over all tiles. To do so, we determine the edge ratios of a tile nm,

$$u_{nm\rightarrow} = \frac{\sum_l \tilde{P}_{((n+1)N_g-1,mN_g+l)}}{\sum_l \tilde{P}_{((n+1)N_g,mN_g+l)}} \tag{2}$$

for the right edge and

$$u_{nm\downarrow} = \frac{\sum_k \tilde{P}_{(nN_g+k,(m+1)N_g-1)}}{\sum_k \tilde{P}_{(nN_g+k,(m+1)N_g)}} \tag{3}$$

for the bottom edge. These edge ratios are modified by applying the tiling correction factors $C_{nm}$, and we find the factors minimizing the deviation of the modified ratios from unity using

$$C_{nm} = \arg\min \sum_{n,m} \left|\log\left(\frac{u_{nm\rightarrow} C_{nm}}{C_{n+1,m}}\right)\right| + \left|\log\left(\frac{u_{nm\downarrow} C_{nm}}{C_{n,m+1}}\right)\right|. \tag{4}$$

The tiling corrected data are then given by $\check{p}_{ij} = \tilde{p}_{ij} C_{nm}$. An example of the tiling correction is shown later in Fig. 5.

## III. RESULTS

### A. Point spread function

To measure the experimental three-dimensional PSF of the setup, a 0.5 $\mu$m diameter PS bead on glass in air (sample C) was imaged. Flat-field corrected images $\tilde{p}_{ij}$ in the $x$-$y$ and $y$-$z$ planes are shown in Fig. 3. The in-plane PSF full width at half maximum (FWHM) is 0.5 $\mu$m, while the axial PSF FWHM is 2.9 $\mu$m, comparable to previous single point scanning data using a 0.75 NA objective.[27]

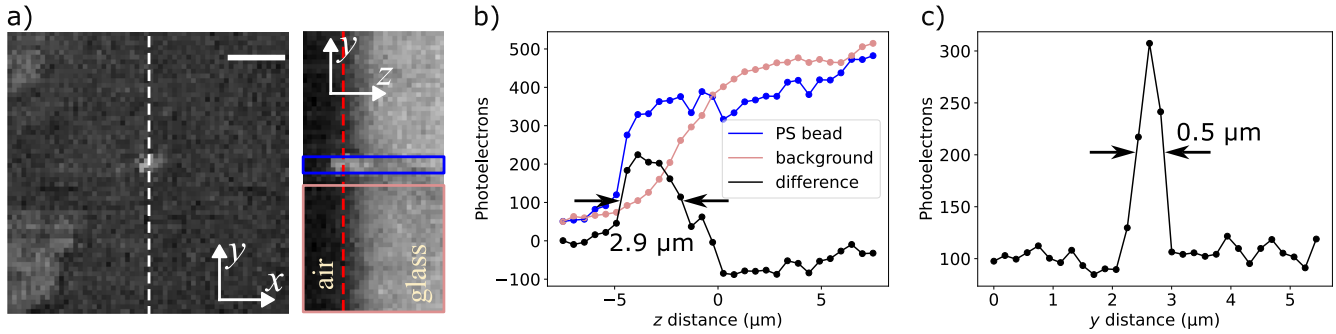


**FIG. 3.** (a) CARS PSF measured on a 0.5 $\mu$m PS bead (sample C) using 30 $z$ planes with 0.5 $\mu$m step. Pump 785 nm, 20 mW and Stokes 50 mW, 3050 cm$^{-1}$. $N_g$ = 40, $S_{int}$ = 10, $g_c$ = 30 dB, and $\nu_f$ = 100 Hz. $\tau_e$ = 8 ms. (a) $xy$ and $yz$ sections, with corresponding positions are indicated by the dashed lines; greyscale 0 to 583 pe; scale bar 2 $\mu$m. (b) z-dependent signal of the glass–air interface without the PS bead [pink, averaged over the pink rectangle in panel (a)], with the PS bead [blue, averaged over blue rectangle in panel (a)], and their difference (black), showing the axial PSF. (c) Signal along the white dashed line in panel (a) showing the lateral PSF.

## B. Tiling correction

To demonstrate the efficacy of the tiling correction on biomedical samples of interest, a human skin sample (see Sec. II A) was imaged at 2935 and 2970 cm$^{-1}$. The resulting data $p_{ij}$ using $S_{int}$ = 10 are shown in Fig. 4(a). The vignetting due to the excitation beam profiles is clearly visible, together with the tiling. After flat-field correction using $r_{nm}$ from sample O and $v = 1/3$ (see Fig. 2), as shown in Fig. 4(b), the resulting $\tilde{p}_{ij}$ given in Fig. 4(c) shows that the vignetting has been corrected, but the tiling, while somewhat reduced, is still clearly visible. The change in the tile pattern indicates that the origin of the remaining pattern is different from the tiling in a homogeneous medium. Applying the tiling correction, Eq. (4) yields the factors $C_{nm}$ shown in Fig. 4(d). The corrected $\check{p}_{ij}$ shown in Fig. 4(e) is void of significant tiling, confirming that the correction is effective. The visible structure is attributed to basal cells (circular objects) in a collagen matrix and elastin (tubular shape).

## C. Three-dimensional bead imaging

To demonstrate the $z$-sectioning capability on strongly scattering structures, sample B hosting a mixture of 20 $\mu$m PMMA beads and 10 $\mu$m PS beads in oil was imaged. A $z$-stack over 10 positions with 2 $\mu$m increments was measured. Using different pump–Stokes delays, CARS stacks at 2930 cm$^{-1}$ dominated by oil, at 2950 cm$^{-1}$ dominated by PMMA, and at 3050 cm$^{-1}$ dominated by PS were sequentially recorded. Flat-field and tiling correction was applied, and the resulting $\check{p}_{ij}$ is shown in Fig. 5. We can see that the method is able to create well-resolved $z$-sectioning also for strongly scattering samples. Comparing this with detection imaging, as previously used[19,22] and shown in Fig. S2, the significant advantage of excitation imaging created by the non-linear exitation PSF is evident. Nevertheless, the 10 $\mu$m PS beads with their large refractive index contrast to olive oil (1.59 to 1.46) provide such large deviations in the detection imaging that the signals from different foci are mixed, as shown in Fig. S3. This mixing leads to the shadowing of the lower PS beads by the beads lying on top of them, as shown in Fig. 5(a). Considering that the shadowing shows the obstacles in detection direction, increasing the NA of excitation and detection would spatially blur and thus reduce the shadowing.

## D. CARS and SHG imaging of human skin

To demonstrate the performance of the multifocal approach on medically relevant scattering samples, a human skin tissue was

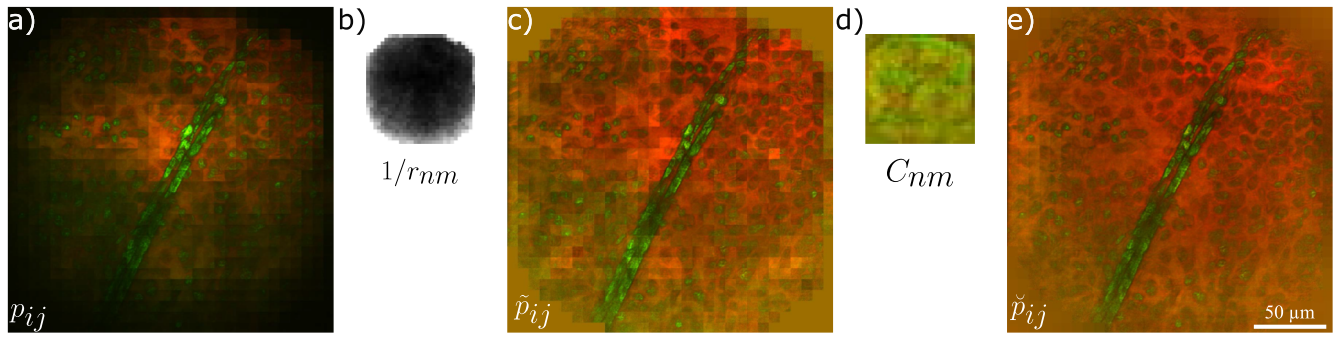


**FIG. 4.** CARS false color RGB image of a human skin sample demonstrating the flat field and tiling correction. $\nu_f$ = 50 Hz, $\tau_e$ = 15 ms, $g_c$ = 20 dB, $N_g$ = 30, and $S_{int}$ = 10. Pump 785 nm, 18 mW; Stokes 48 mW. CARS was imaged at 2935 cm$^{-1}$ (red) and then 2970 cm$^{-1}$ (green) by changing the pump-Stokes delay. (a) Uncorrected $p_{ij}$, range 0 – 10 138 pe for 2935 cm$^{-1}$ and 0 – 5530 pe for 2970 cm$^{-1}$. (b) Flat-field correction factors $1/r_{nm}$, greyscale from 0 (black) to 4 (white). (c) Flat-field corrected data $\tilde{p}_{ij}$, range 0 to 3226 pe for 2935 cm$^{-1}$ and 0 to 5529 pe for 2970 cm$^{-1}$. (d) Tiling correction factors $C_{nm}$ with color coding as (c) on a range from 0 to 2. (e) Tiling corrected $\check{p}_{ij}$; same range as in panel (c).

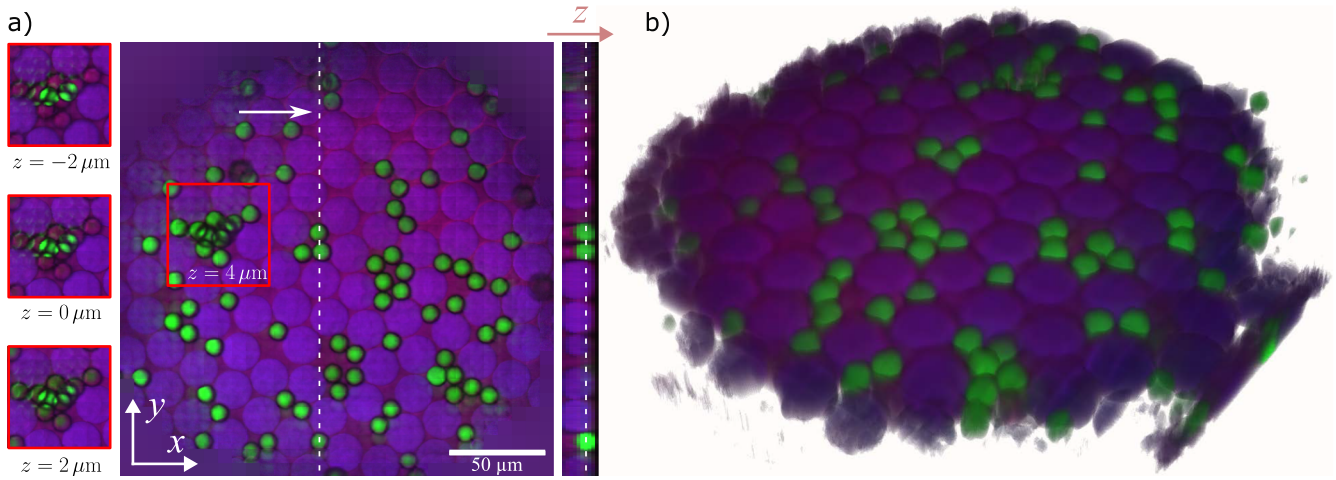


**FIG. 5.** Three-dimensional CARS imaging of sample B at 3050 $cm^{-1}$ (green) highlighting the 10 $\mu$m PS beads, at 2950 $cm^{-1}$ (blue) highlighting the 20 $\mu$m PMMA beads, and at 2930 $cm^{-1}$ (red) showing the oil used as an embedding medium. $\nu_f$ = 100 Hz, $\tau_e$ = 5 ms, $g_c$ = 10 dB, $N_g$ = 30, and $S_{int}$ = 10. Pump 785 nm, 10 mW; Stokes 50 mW. Stack of 10 *z*-steps of 2 $\mu$m size. (a) Images $\check{p}_{ij}$ after flat-field $(v = 1/5)$ and tiling correction using value ranges of 0 to 14 600 pe for 2930 $cm^{-1}$, 0 to 17 520 pe for 3050 $cm^{-1}$, and 0 to 29 142 pe for 2950 $cm^{-1}$. (a) $x - y$ section, with a zoom of an area indicated as the red square at different *z* planes with 2 $\mu$m increments shown on the left and a $y - z$ section shown on the right. The white dashed lines indicate the *x* and *z* positions of the $y - z$ and $x - y$ sections. (b) 3D reconstruction of $\check{p}_{ij}$.

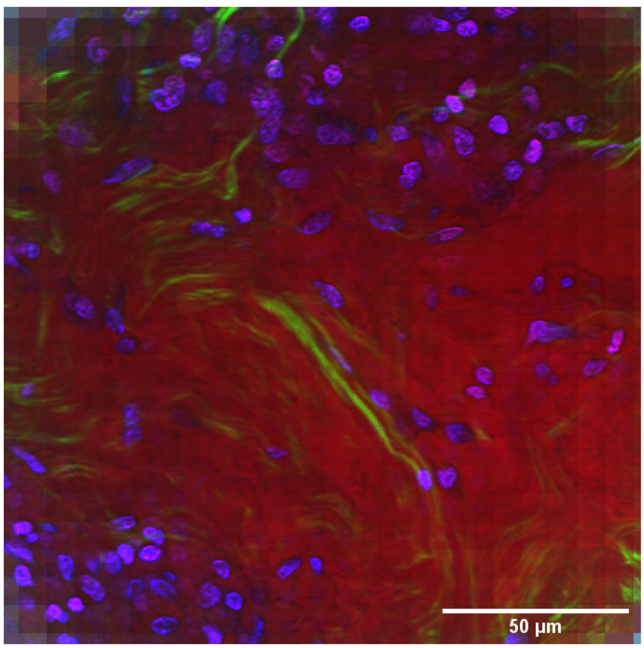


**FIG. 6.** CARS and SHG imaging of human skin tissue. $\nu_f$ = 200 Hz, $\tau_e$ = 4 ms, $g_c$ = 20 dB, $N_g$ = 30, and $S_{int}$ = 10. Pump 785 nm, 20 mW; Stokes 50 mW. Flat-field $(v = 1/5)$ and tiling correction. Blue channel is CARS at 2970 $cm^{-1}$, red channel is CARS at 2935 $cm^{-1}$, and green channel is SHG. Value ranges are 0 to 8294 pe for the red, 0 to 4608 pe for the green, and 0 to 3686 pe for the blue channel. The total acquisition time was 13.5 s.

imaged (see Sec. II A). CARS was recorded at 2935 and 2970 $cm^{-1}$. To measure SHG of the Stokes, the FB650-40 filter was replaced by a BP520-40 (Thorlabs), and the pump was blocked. The glass blocks were not removed for convenience, but we note that the SHG signal would be increased about 4 times by doing so. The resulting image encoding all three datasets as false color is presented in Fig. 6. We attribute the structure to basal cells (blue), elastin fibers (green), and collagen matrix (red).

## IV. CONCLUSIONS

In this work, we have demonstrated multifocal CARS microscopy with 1089 foci, enabled by a high repetition rate amplified oscillator and an optical parametric amplifier. Importantly, different from previous multifocal approaches,[19,22] we introduced employing the camera detection only to separate the signals of the different foci, and not for imaging. This retains the insensitivity to sample scattering afforded by the non-linear excitation PSF created by longer wavelength light, which is the hallmark of point-scanning CRS. Notably, speckle widefield CARS[16,17] instead is using detection imaging and also has intrinsic speckle noise. As with other multifocal techniques,[21] the axial resolution is the same as in single point scanning, since the PSFs of the individual foci are the same. There is crosstalk between foci for large defocus distances, which depends on the density of foci, but the third-order non-linearity in CARS reduces this effect compared to TPF or confocal fluorescence. The influence of scattering and absorption by the sample on the penetration depth for each individual focus is the same as in single point scanning, but multiple foci can result in crosstalk due to

elastic scattering of the CARS signal by the sample in the detection path.

The speed of the technique in the present setup is limited by the camera readout. The laser source caters for up to 100 times higher excitation powers then presently used, up to 1 W for the pump and 5 W for the Stokes. This would provide up to $10^6$ times higher CARS powers, which, detecting the same number of CARS photons per pixel, would allow to decrease the exposure time accordingly, thus providing a signal for up to $10^8$ Hz camera frame rate. The practically usable power will be limited by sample damage, which depends on the samples imaged. Comparing with present single point scanning techniques, which have shown single pixel rates of 10 MHz at laser repetition rates around 50 MHz, we can estimate that using 1000 foci, with 100 times less repetition rate, providing the same energy per pulse per focus, a 10 times faster imaging is possible, likely with lower damage due to the lower repetition rate. This would enable 100 Hz rate for megapixel CARS images. Since the camera is only used to separate the different foci rather than to image the sample at the Nyquist limit, only a small number of pixels are required, for example, 100 × 100 pixels would suffice for 1000 foci, which can be read by present camera technology (for example the IX cameras i-SPEED 5 series or the Pharsighted E9 series cameras) at up to $10^6$ Hz frame rate, opening the perspective of megapixel CARS imaging with 1000 Hz rates. To support such high rates, the focus array shift would need to be done non-mechanically. Importantly, since the shift only needs to cover the small range between the two foci, resolving less than 100 points per direction, electro-optical deflectors would be suited (for example, ConOptics Model 412 providing microsecond response times).

When choosing the number of foci, the main consideration is the available laser pulse energy, which should provide each focus with sufficient pulse energy to efficiently generate CARS. Assuming a $(300\ \text{nm})^2$ focus area and 1 ps pulses, this energy is of the order of 1–10 nJ in typical organic samples, as shown here, scaling with pulse duration and focus area. For a typical 80 MHz oscillator repetition rate, this requires some 100 mW power. Thus, 1000 foci require pulse energies of 1 to 10 $\mu$J, and lower repetition rates are needed to limit the power; for example, at 200 KHz, the required power is 0.2–2 W. A further consideration is the provision of diffraction limited foci over the whole array. Modern microscope objectives can image some $10^6$ to $10^8$ resolved points. This number is limiting the number of foci times the number of resolved points per focus coverage area, in our case, around $1089 \times 30^2$, which is about $10^6$ and thus not a limiting factor. In addition, the detection camera needs to be able to image all foci, and since cameras with $10^6$ to $10^7$ resolved points are readily available, this is also not a limiting factor.

As future improvement of the method, time-multiplexing[28] is a promising approach not only for suppression of crosstalk between foci, as demonstrated for TPF,[28] but also to address different vibrational frequencies via spectral focusing, allowing for a spectrally resolved imaging where different foci address different vibrational resonances, doing away with the sequential spectral tuning.

## SUPPLEMENTARY MATERIAL

The supplementary material contains, in Sec. S1, additional data showing the influence of the detection region size $S_{\text{int}}$ on the excitation image and, in Sec. S2, a comparison between detection and excitation imaging of beads.

## ACKNOWLEDGMENTS

We thank Jonas Berzinš (Light Conversion) for proofreading and commenting on the draft, Skaidra Valiukevičien (Lithuanian University of Health Sciences) for providing human skin sarcoma samples, and Paola Borri for support. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska–Curie Grant Agreement No. 812992.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**D. Gudavičius**: Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (equal); Resources (equal); Software (equal); Validation (lead); Visualization (lead); Writing – original draft (equal); Writing – review & editing (equal). **L. Kontenis**: Methodology (supporting); Software (equal); Writing – review & editing (supporting); Supervision (supporting); Project administration (equal). **W. Langbein**: Conceptualization (lead); Formal analysis (supporting); Funding acquisition (supporting); Methodology (lead); Resources (equal); Project administration (equal); Software (supporting); Supervision (lead); Validation (supporting); Visualization (supporting); Writing – original draft (equal); Writing – review & editing (equal).

## DATA AVAILABILITY

Information about the data created during this research, including how to access it, is available from Cardiff University data archive (http://doi.org/10.17035/d.2024.0313958561).

## REFERENCES

[1] A. Zumbusch, W. Langbein, and P. Borri, "Nonlinear vibrational microscopy applied to lipid biology," Prog. Lipid Res. **52**, 615–632 (2013).
[2] C. Zhang, D. Zhang, and J.-X. Cheng, "Coherent Raman scattering microscopy in biology and medicine," Annu. Rev. Biomed. Eng. **17**, 415–445 (2015).
[3] C. Zhang and J. A. Aldana-Mendoza, "Coherent Raman scattering microscopy for chemical imaging of biological systems," J. Phys.: Photonics **3**, 032002 (2021).
[4] H. Lin and J.-X. Cheng, "Computational coherent Raman scattering imaging: Breaking physical barriers by fusion of advanced instrumentation and data science," eLight **3**, 6 (2023).
[5] A. Zumbusch, G. R. Holtom, and X. S. Xie, "Three-dimensional vibrational imaging by coherent anti-Stokes Raman scattering," Phys. Rev. Lett. **82**, 4142 (1999).
[6] L. Gong, W. Zheng, Y. Ma, and Z. Huang, "Higher-order coherent anti-Stokes Raman scattering microscopy realizes label-free super-resolution vibrational imaging," Nat. Photonics **14**, 115–122 (2019).
[7] 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).

[8]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 (2009).
[9]F. Masia, A. Glen, P. Stephens, P. Borri, and W. Langbein, "Quantitative chemical imaging and unsupervised analysis using hyperspectral coherent anti-Stokes Raman scattering microscopy," Anal. Chem. **85**, 10820–10828 (2013).
[10]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).
[11]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).
[12]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).
[13]Z. Xu, K. Oguchi, Y. Taguchi, S. Takahashi, Y. Sano, T. Mizuguchi, K. Katoh, and Y. Ozeki, "Quantum-enhanced stimulated Raman scattering microscopy in a high-power regime," Opt. Lett. **47**, 5829 (2022).
[14]C. Heinrich, S. Bernet, and M. Ritsch-Marte, "Wide-field coherent anti-Stokes Raman scattering microscopy," Appl. Phys. Lett. **84**, 816–818 (2004).
[15]C. Heinrich, A. Hofer, A. Ritsch, C. Ciardi, S. Bernet, and M. Ritsch-Marte, "Selective imaging of saturated and unsaturated lipids by wide-field CARS-microscopy," Opt. Express **16**, 2699 (2008).
[16]C. Heinrich, A. Hofer, S. Bernet, and M. Ritsch-Marte, "Coherent anti-Stokes Raman scattering microscopy with dynamic speckle illumination," New J. Phys. **10**, 023029 (2008).
[17]E. M. Fantuzzi, S. Heuke, S. Labouesse, D. Gudavičius, R. Bartels, A. Sentenac, and H. Rigneault, "Wide-field coherent anti-Stokes Raman scattering microscopy using random illuminations," Nat. Photonics **17**, 1097–1104 (2023).
[18]J. Oreopoulos, R. Berman, and M. Browne, "Spinning-disk confocal microscopy," in *Quantitative Imaging in Cell Biology* (Elsevier, 2014), pp. 153–175.
[19]J. Bewersdorf, R. Pick, and S. W. Hell, "Multifocal multiphoton microscopy," Opt. Lett. **23**, 655 (1998).
[20]M. Straub and S. W. Hell, "Multifocal multiphoton microscopy: A fast and efficient tool for 3-D fluorescence imaging," Bioimaging **6**, 177–185 (1998).
[21]A. Egner, V. Andresen, and S. W. Hell, "Comparison of the axial resolution of practical Nipkow-disk confocal fluorescence microscopy with that of multifocal multiphoton microscopy: Theory and experiment," J. Microsc. **206**, 24–32 (2002).
[22]T. Minamikawa, M. Hashimoto, K. Fujita, S. Kawata, and T. Araki, "Multifocus excitation coherent anti-Stokes Raman scattering (CARS) microscopy and its applications for real-time imaging," Opt. Express **17**, 9526 (2009).
[23]T. Minamikawa, T. Araki, M. Hashimoto, and H. Niioka, "Real-time imaging of laser-induced membrane disruption of a living cell observed with multifocus coherent anti-Stokes Raman scattering microscopy," J. Biomed. Opt. **16**, 021111 (2011).
[24]R. D. Cardiff, C. H. Miller, and R. J. Munn, "Manual hematoxylin and eosin staining of mouse tissue sections," Cold Spring Harbor Protoc. **2014**, 655.
[25]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).
[26]D. Gudavičius, L. Kontenis, and W. Langbein, "Dual coherent anti-Stokes Raman scattering microscopy across the vibrational spectrum enabled by single pump optically synchronised oscillators and spectral focusing," J. Raman Spectrosc. **55**, 744 (2024).
[27]A. Karuna, F. Masia, P. Borri, and W. Langbein, "Hyperspectral volumetric coherent anti-Stokes Raman scattering microscopy: Quantitative volume determination and nacl as non-resonant standard," J. Raman Spectrosc. **47**, 1167–1173 (2016).
[28]V. Andresen, A. Egner, and S. W. Hell, "Time-multiplexed multifocal multiphoton microscope," Opt. Lett. **26**, 75 (2001).