A novel method of obtaining high-quality Raman spectra of luminescent samples was tested using cyclohexane solutions which had been treated with a fluorescent dye. The method involves removing the fixed pattern irregularity found in the spectra taken with CCD detectors by subtracting spectra taken at different, closely spaced spectrometer positions. It is conceptually similar to SERDS (shifted excitation Raman difference spectroscopy) but has the distinct experimental advantage that it does not require a tunable laser source. The subtracted spectra are obtained as the raw data are converted into a more recognisable and conventional form by iterative fitting of appropriate double Lorentzian functions whose peak parameters are then used to ‘reconstruct’ a conventional representation of the spectrum. Importantly, it is shown that the degree of uncertainty in the resultant ‘reconstructed’ spectra can be gauged reliably by comparing reconstructed spectra obtained at two different spectrometer shifts (δ and 2δ). The method was illustrated and validated using a solvent (cyclohexane) the spectrum of which is well known and which contains both regions with complex overlapping bands and regions with isolated bands. Possible sources of error are discussed and it is shown that, provided the degree of uncertainty in the data is correctly characterised, it is completely valid to draw conclusions about the spectra of the sample on the basis of the reconstructed data. The acronym SSRS (subtracted shifted Raman spectroscopy; pronounced scissors) is proposed for this method, to distinguish it from the SERDS technique.

Keywords: Raman spectroscopy; resonance Raman spectroscopy; luminescence suppression; background subtraction; charge coupled device detectors; shifted excitation; Raman difference

Raman spectroscopy is currently enjoying a strong resurgence in interest with increasingly frequent reports of its application outside research laboratories in harsh industrial environments. Much of this growth in the application of the technique has been due to falling cost of laser sources, increasingly robust and efficient detectors and the availability of holographic filters used for rejection of Rayleigh-scattered light. However, one of the factors which prevents the application of the technique to a broad range of sample types (both industrial and academic) is the occurrence of strong luminescence which can, in many cases, dominate the signals recorded from the sample. This luminescence can arise either from the analyte of interest (if it absorbs at the laser wavelength used for excitation, as in resonance Raman spectroscopy) or from adventitious impurities within unpurified samples. In many cases the background luminescence is broad and featureless so that it is relatively straightforward to remove even a strongly sloped background using curve fitting or Fourier filtering methods. However, removal of a sloping background from Raman spectra can only be a cosmetic exercise since it does not help with the second, and more fundamental, problem associated with intense background signals, namely the high level of photon shot noise the spectra contain.

The problem of sample luminescence has been recognised for decades and many different strategies have been developed to circumvent the problem. It may be possible to quench the luminescence (most commonly by using surface-enhanced Raman techniques); to shift the excitation wavelength to one where the Raman signal lies in a different wavelength range to the luminescence or where the excitation source does not generate the luminescence to use time-domain techniques based on gated or frequency domain detectors which take advantage of any difference in the temporal characteristics of the Raman and background signals, or to monitor the Raman signals at the anti-Stokes side of the excitation wavelength. Many of these strategies have been highly successful for particular sample types but none can be regarded as universal.

An alternative approach, which could not be made practical until recently, is to take advantage of high quantum efficiency (QE) of multichannel detectors, coupled with high throughput spectrographs, to accumulate total signals sufficiently large that the shot noise associated with the background is insignificant compared with the Raman signal levels. The total accumulation times needed to reach this level can be surprisingly short. For example, if a Raman signal of 100 detected photons per second (phtn s⁻¹) is superimposed on a very intense broad luminescence background which is 100 times larger than the Raman signal (i.e., 10⁴ phtn s⁻¹) then the Raman band is barely detectable after 1 s of signal accumulation since the photon shot noise is √1⁰⁰ = 100 phtn, which is the same size as the signal. However, after just 100 s of accumulation the signal will be 10⁴ phtn while the shot noise is 10² phtn, i.e., 10 times lower than the Raman signal. Essentially, the photon shot noise associated with the background decreases as the square root of signal accumulation time and, as the example above shows, it can be decreased to acceptable levels within reasonable accumulation times (even if the relative magnitudes of signal and background appear unfavourable), provided that the total number of detected photons is sufficiently high. CCD detectors and single stage spectrographs give these high detected photon levels so that if they are used the main problem would therefore appear to reduce to the cosmic exercise of subtraction of a smooth background function to generate a Raman signal which lies well above the noise level.

Unfortunately, CCD detectors suffer from an inherent problem in that each detector element (or column of binned elements) has a slightly different response, typically ±1%, so that an evenly illuminated CCD gives a spectrum with an apparently random noise level of approximately 1% of the total signal height. This means that if a Raman signal with a large background is recorded the apparent noise level in the spectrum will never fall below ∆1% of the background. In the example quoted above, although the signal which was accumulated for
100 s would have a shot noise of $10^3$ phton and a Raman signal of $10^4$ phton, the 1% irregularity would be 1% of the background level of $10^6$ phton, i.e., $10^4$ phton, which is the same height as the signal. This 1% irregularity will remain irrespective of the accumulation time used; increasing the total accumulation time will reduce the photon shot noise on the spectrum, but this is not the dominant noise source so that accumulating spectra for longer and longer times is pointless. This means that although the CCD/high throughput detector reduces one noise source (the photon shot noise) because of its capacity for efficient photon detection, it then introduces a second noise source in its place. In principle, it should be possible to flat-field the detector, i.e., to normalise its response, but the method has been found to be ineffective because the intensity distribution of the Raman signal which falls on the detector is not uniform. Alternatively, it is possible to construct scanning multichannel instruments, in which the spectrum is moved along the active area of the detector, but this is a considerable undertaking.

The most successful approach to removing the fixed-pattern non-uniformity (short of constructing a completely new instrument) is to generate a second signal which also contains the detector irregularity and then to subtract this signal. The background can be generated from a simple fluorophore which has no fluorescent bands, provided that the spectrometer grating of the spectrometer rather than the laser. This method takes advantage of a tunable laser source. Here we report a conceptually similar sub-Routine which allows interactive, iterative fitting of spectroscopic data by conventional Gaussian or Lorentzian functions. The software contains a least-squares curve-fitting sub-routine which allows interactive fitting of Raman bands on broad luminescence backgrounds, and appears to constitute a simple and straightforward approach to the whole problem of analysing luminescent samples by Raman methods. It does not require any additional specialised equipment over conventional single wavelength excitation/CCD detector systems and as such allows studies to be carried out using existing non-tunable laser sources which may have wavelengths for which no tunable laser sources exist.

Here we report studies on a conventional model system, cyclohexane solution treated with an organic dye to provide the fluorescence background, which is used to illustrate and to validate the general approach.

### Experimental

Raman spectra were recorded using 363.8 nm excitation (100 mW) from a Spectra-Physics (San Jose, CA, USA) Model 2020 argon ion laser with a 180° back-scattering geometry and a Kaiser (Ann Arbor, MI, USA) holographic notch filter. Scattered photons were collected, dispersed by a Jobin-Yvon (Longjumeau, France) HR640 single stage spectrograph and detected with a Princeton Instruments CCD detector (Princeton, NJ, USA) (PI LN/1152UV). The spectrometer was calibrated using the standard Raman band positions of solid naphthalene.

For difference spectroscopy, the spectrometer grating was moved manually from its initial calibrated position to the required shifted value (δ nm away*) by monitoring the position of a strong line from a medium pressure Pen-Ray Hg lamp (Oriel, Stratford, CT, USA) in real time. The normal acquisition protocol was to record a spectrum at the initial position, shift the spectrometer by δ cm⁻¹ and record a second spectrum, then to move to the third position (approximately 2 δ cm⁻¹ from the original position) for the final data acquisition of the cycle. In order to minimise the effect of changes in background luminescence level due to sample decomposition/laser damage, this three step acquisition cycle was normally repeated several times to average out, as far as possible, gradual changes in sample luminescence or excitation laser power.

The standard test solution was made by dissolving 1,3,5-tri-phenyl-4,5-dihydropyrazole in cyclohexane (approximately 5 × 10⁻⁵ mol dm⁻³). There were clear indications of gradual sample decomposition (loss of luminescence background) during each acquisition cycle so fresh aliquots of the sample were used for every data set.

Raman data were transferred to Grams/386 software for processing. Under the experimental conditions used, the Raman bands of simple test compounds approximated well to simple Lorentzians. The software contains a least-squares curve-fitting sub-routine which allows interactive fitting of spectroscopic data by conventional Gaussian or Lorentzian functions. This sub-routine was modified to include interactive fitting of double Lorentzian bands of the type

$$I = \frac{Ho^2}{\sigma^2 + (v - \nu)}} - \frac{F}{\sigma^2 + (v - v) - \delta^2}$$

The function is composed of a pair of Lorentzian bands, one positive and one negative, both of which have height, width and centre position H, F and v, respectively; δ gives the separation between the peak centres and $F$ is an adjustable scaling factor which allows unsymmetrical double Lorentzians to be treated. In the fitting process the line widths (σ) were constrained to be ≥5.4 cm⁻¹, which is slightly narrower than the instrument-limited width of about 7 cm⁻¹.

### Results

Fig. 1 shows the Raman spectra of dyed cyclohexane solutions which were recorded at three different spectrometer positions about 20 cm⁻¹ apart. The only Raman features which are visible on top of the broad and intense fluorescence background in any of the spectra are the strongest C–H bands around 2900 cm⁻¹. The inset shows an expansion of the data in the C–H region. If a single spectrum is inspected there appears to be significant ‘random’ noise on the spectra but comparison of the shifted spectra shows that most of this ‘noise’, although randomly

---

* Shifting the spectrometer’s central wavelength position shifts all the spectral features by the same number of detector pixels. Since the dispersion of the spectrometer is linear in nm, all the features are moved by the same nm increment. This means that the shift in the position of the Raman bands in cm⁻¹ is different across the spectrometer range. In the example given, the shift δ = 0–3 nm corresponds to an average shift of about 22 cm⁻¹ but the values vary from about 21 to 23 cm⁻¹ across the spectrometer range. Throughout the text the δ values are given in approximate cm⁻¹ units, for clarity, but the fitting procedure was carried out with the correct shift values.
distributed in each spectrum, is reproduced between the spectra.†

The difference spectrum, obtained by subtracting the data in Fig. 1(a) from that in Fig. 1(b) (i.e., a shift of about 20 cm⁻¹) is shown in Fig. 2(a). Superimposed on the difference spectrum is a polynomial baseline function. Fig. 2(b) shows the same data after subtraction of the baseline function; also shown [Fig. 2(c)] is the baseline-corrected difference spectrum for the approximately 40 cm⁻¹ shift.

Fig. 3 compares the difference spectrum with \( \delta = 20 \text{ cm}^{-1} \) [Fig. 2(a)] and the curve-fit to the data using the double Lorentzian function given under Experimental. Fig. 3(c) shows the residual of the curve fit. Fig. 3(d) and (e) are reconstructed conventional spectra for the approximately 20 and 40 cm⁻¹ shifts, which were prepared by generating single Lorentzian peaks using the height, width and cm⁻¹ position data obtained from the double Lorentzian curve fitting process.

Fig. 4 compares the ‘best guess’ reconstructed spectrum with the conventional Raman spectrum of cyclohexane, which was recorded with the same spectrometer under identical experimental conditions, apart from the absence of the fluorescent dye.

**Discussion**

The data in Figs. 1 and 2 clearly demonstrate that the shifted-spectrometer method provides a means of minimising CCD fixed pattern non-uniformity, which is the major noise source in the Raman spectra of luminescent samples recorded with such detectors. The fact that both the Raman and luminescence background signals are shifted together means that, in contrast to the SERDS method, some residual background signal must always remain after the difference spectrum is generated. However, this is the only additional complication which arises from the use of this more experimentally general method. This residual background signal in the difference spectra must be removed before the derivative-like Raman bands can be curve fitted.

The problem of background subtraction is rather less severe than one might imagine since, if the background can be represented a polynomial of the form \( y = a + bx + cx^2 + dx^3 + \ldots \), then in the difference spectra the zero-order terms disappear, the first-order terms give a constant offset in the difference spectra, the second-order terms give an easily subtracted linear baseline and it is only the third- and higher order terms which give a non-linear baseline in the difference spectra. It is clear from Fig. 1 that unscaled subtraction of the spectra obtained at the three spectrometer settings will give difference spectra where the derivative-like Raman bands sit on a background which falls smoothly to a minimum at approximately the centre of the spectral range covered.† It is possible to remove this smooth background from unscaled difference spectra (i.e., spectra generated by simple subtraction of the spectra at the different spectrometer positions), but we chose to scale the subtraction so that the average background level in the resultant difference spectra is closer to zero. This choice makes polynomial line fitting (and then removal) of the residual luminescence background easier and also means that the absolute magnitude of the fixed pattern noise on the background is as close as possible in both spectra, and will therefore be most efficiently cancelled in the difference spectra. Provided that the scaling factor used in the subtraction is known it can be included in the double Lorentzian function used for curve fitting. In practice, we have found that the scaling factor (\( F \)) is normally

![Fig. 1](image1.png)  
**Fig. 1** Raman spectra of dye-treated cyclohexane at three different spectrometer positions, (a)–(c), separated by approximately 20 cm⁻¹. Inset: enlarged views of the spectra in the C–H stretching region which have been vertically offset for clarity. \( h_{\text{ex}} = 363.8 \text{ nm} \); laser power at sample = 100 mW; total accumulation time for each spectrum = 25 min.

† To correct for gradual drift in signal level between the data taken at each of the three spectrometer positions, it is useful to offset, digitally, the spectra by the same amount that they were shifted by (in effect cancelling the effect of the shift in the spectrometer) and check that the spectra overlay exactly. If they do not, then the signals at each of the three spectrometer positions can be normalised before further processing is carried out.

![Fig. 2](image2.png)  
**Fig. 2** Difference Raman spectrum of the dyed cyclohexane test solution generated by scaled subtraction of the spectra shown in Fig. 1(a) and (b), i.e., the difference spectrum given by an approximately 20 cm⁻¹ shift in spectrometer position. Superimposed on the spectrum is the polynomial line used for background subtraction. (b) Difference Raman spectrum at approximately 20 cm⁻¹ shift after subtraction of the background function. (c) Difference Raman spectrum of the same sample with an approximately 40 cm⁻¹ shift, after subtraction of a background function.

† It is possible to judge how much curvature will appear in the difference spectra simply by differentiating a single spectrum without actually recording spectra at a series of different spectrometer positions. In such a differentiated spectrum the noise level is much higher than in the differenced spectra because of the fixed-pattern detector irregularity, but the relative magnitudes of the Raman bands compared with the residual background signal follow the same pattern as is seen in the shifted-difference spectra.
between 0.95 and 1.0 so that even the corrected fitting function appears close to a simple double Lorentzian.

Once the difference spectra have been generated and their backgrounds removed (Fig. 2), it only remains to fit the curves with the double Lorentzian functions given above. This generates fits which can be displayed graphically (Fig. 3) or tabulated as a series of peak heights, widths and positions. Although this table will contain all the data collected in the experiment, the data are more easily interpreted and compared with other Raman spectra if the parameters are used to reconstruct a conventional spectrum composed of the simple Lorentzian peaks (see Fig. 3). Although this final step is useful it can also be deceptive, since it generates spectra with no random noise. Such data are rightly regarded with suspicion since they are counter-intuitive. The root of the problem is that we are accustomed to judging the quality of the data (i.e., the degree to which they could be reproduced) in a given spectrum by comparing the peaks with the random noise. With no random noise is not possible to make this judgement. With test data, where the peak positions are already known, there is no problem but with spectra of real samples, where there is no independent method of validation, the issue of confidence level on the basis of the quality of the curve fit to the difference data, but this is less than direct. It was for this reason that we adopted the strategy of taking spectra at three different, approximately evenly spaced, grating positions. Subtraction of these spectra gives two difference spectra, which have shifts of approximately δ and 2δ. These difference spectra are different from each other because those bands which partially cancel their intensities at one shift will not also cancel at the other. Independently background correcting and curve-fitting both of these spectra to appropriate double Lorentzian functions generates two peak tables, which can then be used to construct two conventional Raman spectra, as shown in Fig. 3(d) and (e). These spectra can then be compared to give an indication of the degree of reproducibility of the peaks. It would not be expected that the reconstructed spectra would be identical with each other but it is the level of irreproducibility which is the critical factor.

As a final step, it is useful to generate a single ‘best guess’ spectrum. To do this, the peak data from the different shifted spectra are combined to give a new set of parameters, which can then be used as the initial guess for the least squares fitting. Average values are taken for the peaks which appear in both spectra while those peaks which appear in only one of the spectra (and which can arise because of the over-enthusiastic curve fitting to the difference data) are eliminated. Iterative fitting of this input parameter set to both the difference spectra yields two new sets of peak parameters which can then be used to reconstruct a pair of nearly identical conventional Raman spectra. The average of these two spectra corresponds to the ‘best guess’ spectrum.

The success of this overall approach is demonstrated by the data for cyclohexane given in Figs. 1–4. Cyclohexane was chosen because it has well defined and separated Raman bands in the region 1000–1500 cm\(^{-1}\), which illustrate the double Lorentzian form of the difference spectra and are very straightforward to curve-fit. In addition, it also has a complex band shape in the C–H region at approximately 2900 cm\(^{-1}\), which is a much more stringent test of whether the fitting method can be applied with confidence to non-ideal spectra, where there is considerable overlap between Raman bands.

There are two separate points to address: the first is the degree to which the reconstructed data reproduce the actual known spectrum of the pure solvent; the second is whether it would be possible to determine the degree of uncertainty in the reconstructed spectra, even if there were no independent data to provide verification. In the original data even the strongest C–H Raman bands are barely visible above the fixed-pattern noise of the detector; indeed, the lower frequency bands at approximately 1000–1400 cm\(^{-1}\) are almost as well defined, despite their much lower absolute intensity, because the luminescence intensity is also lower at this end of the spectrum. Nonetheless, in both of these spectral regions the largest bands in the ‘best fit’ reconstructed spectra are in excellent agreement with those of the Raman spectrum of pure cyclohexane recorded using the same spectrometer (enlarged views of the data in the C–H region are given in Fig. 5). Moreover, the method has also given

![Fig. 3](image1.png) (a) and (b) overlaid views of the approximately 20 cm\(^{-1}\) shift difference spectrum [from Fig. 2(b)] and a double Lorentzian curve-fit to this spectrum. The residual from this curve-fit is shown on the same vertical axis as (c). (d) and (e) conventional Raman spectra reconstructed from the curve-fit data obtained with spectrometer shifts of approximately 20 and 40 cm\(^{-1}\), respectively. Inset: enlarged views of a small section of the reconstructed spectra at both wavenumber shifts, which have been overlaid for comparison.

![Fig. 4](image2.png) (a) ‘Best guess’ reconstructed Raman spectrum of the dyed cyclohexane sample. (b) Conventional Raman spectrum of undyed cyclohexane.
good data for the positions and intensities of the much weaker cluster of C–H bands which lie at approximately 2700 cm\(^{-1}\) and which were completely invisible in the original (unsubtracted) spectra. Even the strongest band in this cluster is 18 times weaker than the strongest C–H bands.

The second point to address is whether it is possible to judge the confidence which can be placed in the data obtained from the curve-fitting procedure by comparing the reconstructed \(\delta\) and \(2\delta\) spectra. In general, for spectral regions where there is good agreement between the reconstructed \(\delta\) and \(2\delta\) spectra (i.e., where there is little discernible uncertainty in the data), we would expect a correspondingly close agreement between the ‘best fit’ and real spectra. This expectation is borne out by the data. For example, the two strong bands at 2938 and 2852 cm\(^{-1}\) are nearly identical in the reconstructed \(\delta\) and \(2\delta\) spectra, so that a high degree of confidence can be placed in them. This confidence is borne out by the very close agreement of these bands in the ‘best fit’ and real spectra. On the other hand, less confidence can be placed in the absolute intensity of the band at 2924 cm\(^{-1}\), since the reconstructed spectra at different shifts disagree to some extent in intensity, although not in the width or band centre. The intensity of this band in the ‘best guess’ spectrum is very close to the actual intensity, but this does not remove the fact that it does have some uncertainty associated with it.

It is only possible to detect differences between the authentic and reconstructed spectrum by looking carefully at the weaker bands. Careful inspection shows that, for example, the relative intensity of the two small bands around 2890 cm\(^{-1}\) in the spectrum of cyclohexane is not correctly reproduced in the ‘best fit’ reconstructed spectrum. In addition, the shoulder on the weak cyclohexane band at about 2700 cm\(^{-1}\) does not appear in the ‘best fit’ spectrum. It is instructive to consider the sources of inaccuracy in these two regions separately as well as determining whether the low confidence that should be placed in the data in these regions could be determined even if an authentic spectrum were not available for comparison.

The region where there is the worst agreement between the ‘best guess’ and actual spectra is around 2890 cm\(^{-1}\), which is not surprising since this region contains weak bands which overlap with much stronger features on either side. In the difference data the negative components of the stronger bands will tend to obscure their weaker neighbours. Again, this is the region where there is the worst agreement between reconstructed \(\delta\) and \(2\delta\) spectra so that it would be clear that this was a region where little confidence could be placed in the position and absolute intensities of the weaker peaks in the ‘best guess’ spectrum, even if the actual spectrum were not known. The main source of the uncertainty in the data in this region is not actually a low signal-to-noise ratio but arises from the need to fit multiple overlapping double Lorentzians to the difference spectra. The curve-fits shown in Figs. 3–5 were based on six such double Lorentzians in the region 2800–3000 cm\(^{-1}\), but we have found that we need seven Lorentzian functions to fit even the simple Raman spectrum of cyclohexane in this 200 cm\(^{-1}\) spectral range, so that it is not surprising that the fit in this range is acceptable rather than faultless. After finding that seven Lorentzians were needed to fit the actual spectrum of cyclohexane, we could have improved the quality of the fit to the difference data by adding another double Lorentzian during the curve-fitting process, but we deliberately refrained from doing so, since we were specifically interested in what spectra could be extracted from the difference data in the case where no authentic spectra of the sample were available to guide the fitting process.

The other source of possibly inaccuracy in the fitted spectra is the actual noise level on the difference data, which is mostly photon shot noise. This random noise is, of course, much smaller than the detector fixed pattern noise but it is not zero and, because it is truly random, it can result in features which appear double Lorentzian in the difference data and will be included in the curve-fit. However, such features are unlikely to occur in difference data taken at two different spectrometer shifts, so that by comparing the reconstructed spectra (as shown in the inset to Fig. 3) it is possible to judge the level of uncertainty due to photon shot noise. Any features which are of the same magnitude as the noise peaks and which appear in only one of the reconstructed spectra should not be considered in any way reliable. Similarly, the method is unlikely to detect real peaks which are around this magnitude because they are of similar size to the effective noise limit. In the example here the effective noise limit is similar to the height of the weak shoulder on the band at approximately 2700 cm\(^{-1}\), which explains why it is not found in the ‘best guess’ reconstructed spectrum. Again, the important point is that the level of uncertainty in the data is known so that the presence or absence of this band in the fitted spectrum is known to be undefined in the data.

Finally, it is important to emphasise that the errors introduced in the fitting method are very small and that the overall agreement between the fitted and actual spectra is excellent.

**Conclusions**

This simple method of taking spectra at several different spectrometer positions removes the fixed pattern irregularity found in the Raman spectra of luminescent samples taken with CCD detectors. The method is conceptually similar to SERDS but has the distinct experimental advantage that it does not require a tunable laser source. We propose the acronym SSRS (subtracted shifted Raman spectroscopy; pronounced scissors) for this method, to distinguish it from SERDS. The main disadvantage of the method is that the subtracted spectra obtained as the raw data need further processing to convert them.

**Fig. 5** (a) and (b) enlarged views of the reconstructed Raman spectra of the dyed cyclohexane sample obtained at shifts of approximately 20 and 40 cm\(^{-1}\). (c) and (d) enlarged views of the ‘best guess’ reconstructed spectrum of the dyed cyclohexane sample and the authentic spectrum of cyclohexane.
into a recognisable conventional form (although they could be utilised in analytical applications directly). However, the fitting process is straightforward and, more important, the degree of uncertainty in the resultant ‘best guess’ spectrum can be gauged by comparing the reconstructed $\delta$ and $2\delta$ spectra. Provided this is done it is completely valid to draw conclusions about the spectra of the sample on the basis of the reconstructed data since the uncertainty levels are low and are quantified.

E. Bourguignon is pleased to acknowledge the receipt of an ERASMUS studentship. The authors thank Professor A. P. de Silva (Queens University, Belfast) for the generous gift of a sample of 1,3,5-triphenyl-4,5-dihydropyrazole.

References