

Smoothing Raman spectra with contiguous single-channel fitting of Voigt distributions - an automated, high quality procedure

Supporting Information

H. Georg Schulze,¹ Shreyas Rangan,¹ James M. Piret,^{1,2} Michael W. Blades³,
and Robin F. B. Turner,^{1,3,4*}

1. Michael Smith Laboratories, The University of British Columbia,
2185 East Mall, Vancouver, BC, Canada, V6T 1Z4.
2. Department of Chemical and Biological Engineering, The University of British
Columbia, 2360 East Mall, Vancouver, BC, Canada, V6T 1Z3.
3. Department of Chemistry, The University of British Columbia,
2036 Main Mall, Vancouver, BC, Canada, V6T 1Z1.
4. Department of Electrical and Computer Engineering, The University of British
Columbia, 2332 Main Mall, Vancouver, BC, Canada, V6T 1Z4.

*Corresponding Author (turner@mssl.ubc.ca)

Here we present information about peak distortions exhibited by the various smoothing algorithms on a synthetic Lorentzian peak as well as smoothing using the Fourier transform (FFT Method).¹ The results are reported in Table SI and the statistical significance tests in Table SII. The Basic Method performed significantly better than the Spline Method, the iSG Method, and the FFT Method on a RMSE basis. The Basic Method performed significantly better than the Spline Method and the iSG Method with regard to peak height, but the FFT Method produced better peak heights than the Basic Method (*cf.* Table SI), though not significantly so (*cf.* Table SII). Regarding peak area, only the FFT Method produced a significantly poorer result.

Though the FFT Method produced good results in terms of peak height and peak position on a single Lorentzian peak (*cf.* Table SI), its performance on a collection of peaks with different attributes, such as present in measured Raman spectra, was unsatisfactory with many peak heights poorly estimated and ringing evident in some sections. An example, using 80 frequencies, is shown in Figure S1. A further disadvantage compared to the fitting methods (e.g. the Basic Method) is that full automation might be difficult to achieve.

Table SI

Smoothing results (mean \pm standard error) for different smoothing methods obtained from 20 simulated Lorentzian bands with 1 % added white noise. The basic Method used a 25 degrees of freedom Voigt distribution and a 10 channel initial and 5 channel intermediate distribution width limit. The fast Fourier transform (FFT) method used the 40 largest amplitude frequencies (the best case in the range 20 to 50 frequencies).
(a.u. = arbitrary units).

Method	RMSE	Position (channel)	Height (a.u.)	Area (a.u. x channel)
Basic Method	0.0080 \pm 0.0003	75 \pm 0	1.4921 \pm 0.0062	22.1064 \pm 0.0403
Spline Method	0.0123 \pm 0.0003	75 \pm 0	1.4316 \pm 0.0025	22.0762 \pm 0.0329
iSG Method	0.0156 \pm 0.0005	75 \pm 0	1.4039 \pm 0.0042	22.0982 \pm 0.0312
FFT Method	0.0110 \pm 0.0003	75 \pm 0	1.5 \pm 0	22.3104 \pm 0.0482
Simulation	0.0 \pm 0.0	75 \pm 0	1.5 \pm 0	22.1209 \pm 0

Table SII

Results of two-tailed *t*-tests performed on the outcomes of the comparisons reported in Table SI. The conservative Bonferroni significance level for 9 tests is $\alpha = 0.05/9$ or 0.0056. The Basic Method (*) was compared to the other methods and the *p*-value is shown for the method against which it was compared.

Method	RMSE	Position (channel)	Height (a.u.)	Area (a.u. x channel)
Basic Method	*	*	*	*
Spline Method	$p < 0.0001$	-	$p < 0.0001$	$p > 0.0056$
iSG Method	$p < 0.0001$	-	$p < 0.0001$	$p > 0.0056$
FFT Method	$p < 0.0001$	-	$p > 0.0056$	$p < 0.0056$
Simulation	-	-	-	-

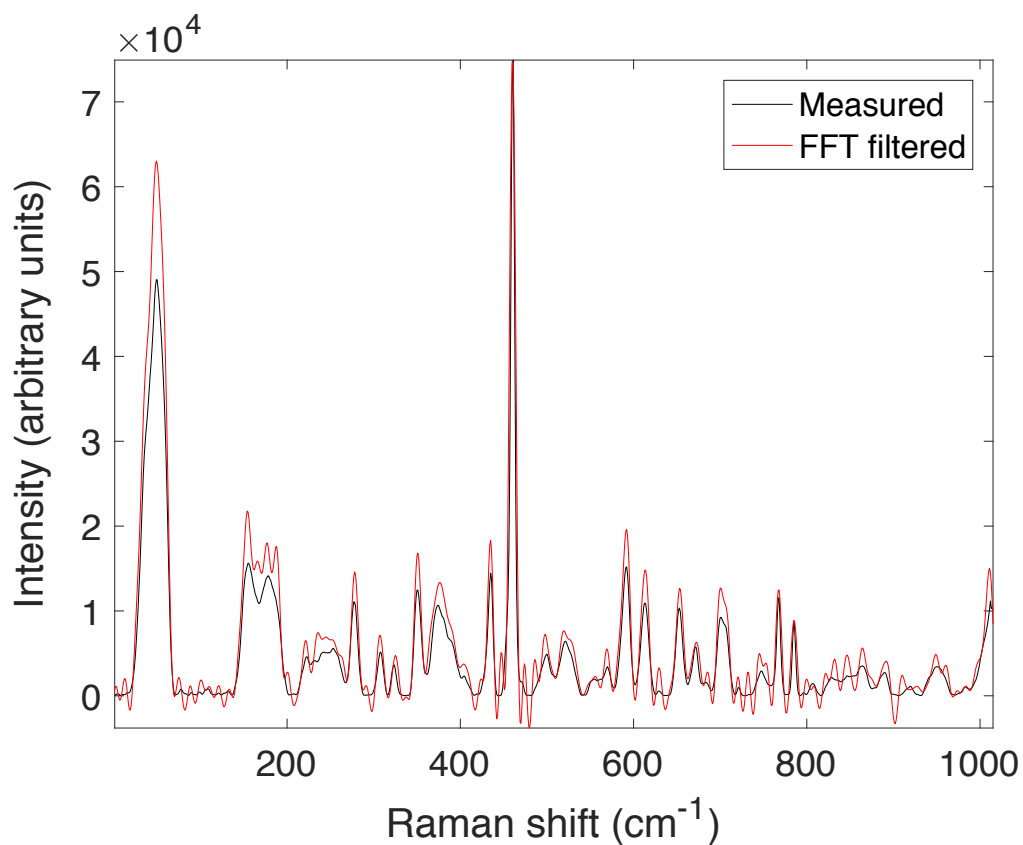


Figure S1

Measured Raman spectrum of formalin treated Chinese hamster ovary cells and filtered with a fast Fourier transform (using 80 frequencies) to reject high frequency noise produced unsatisfactory results.

REFERENCES

1. <https://www.mathworks.com/matlabcentral/fileexchange/25017-fft-filter-clean-your-signals-and-display-results-> (Accessed 26 June 2018).