Novel Methods for Calibration in Raman Spectroscopy

DONGYUE LIU

Supervised by Dr. Bryan Hennelly

Maynooth University
National University of Ireland Maynooth

A thesis presented for the degree of Doctor of Philosophy

DEPARTMENT OF ELECTRONIC ENGINEERING

MAYNOOTH UNIVERSITY

October 2022

Declaration

I hereby certify that the material contained within this thesis is entirely my own work, does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed:

ID Number: 14183684

Date: 10/30/2022

Acknowledgements

First and foremost I am extremely grateful to my supervisor, Dr. Bryan Hennelly, a responsible and resourceful scholar. He always takes lots of time to guide me to think what I should do. Without his valuable instructions, patient guidance and inspiring encouragements, I would not have completed my PhD successfully.

My sincere appreciation also goes to Sinead Barton and Laura Kerr. They helped me to record spectra in our lab and explain to me the questions I did not understand. Also I want to express my gratitude to the staff in Maynooth University helped me every day. An additional thank you to all my friends. Thank you for helping me and supporting me throughout my PhD time.

I will be forever in debt to my family for the kindness, support, and encouragement that in these years.

And finally, a massive thank you to the most important person - Sheng Yang. I never would have finished this PhD without you.

Contents

List of Figures

List of Tables

Novel Methods for Calibration in Raman Spectroscopy

DONGYUE LIU

Abstract

Raman spectroscopy can probe the chemical structure of a material providing an optical 'fingerprint' unique to the sample. Such is the capacity of Raman spectroscopy to identify different materials, it be to classify biological cells and tissue and can provide an 'optical biopsy' for various types of disease. A key component in Raman diagnostics is the use of multivariate statistical algorithms that can be trained using datasets of known samples to classify the groups based on the subtle differences between them. Despite the great progress in this field in recent decades, Raman spectroscopy has

never been adopted clinically. The key reason for this is the poor resproducibility of Raman spectroscopy across instruments; in other words the same material can produce different spectra when recorded using different spectrometers. These differences can include small movement of the Raman peaks along the wavenumber axis (wavenumber miscalibration) or modulation in the amplitude of the peaks (intensity calibration). Such changes can render a multivariate classifier trained on one instrument to be completely useless in identifying samples recorded from another instrument. The overall goal of this thesis is to develop new methods for wavenumber and intensity calibration that can help Raman spectroscopy penetrate into the clinic.

Chapter 1

Introduction

1.1 Introduction

Spectroscopy relates to the recording of the wavelength (or wavenumber) distributing of a light field and has many applications in the identification and classification of different substances, which either absorb or produce spectral irradiance patterns that contain information on the chemical composition. In order to ensure the accuracy of the data collected from a spectrometer, it is essential to calibrate the instrument before the experiment. Even small errors in the position of the components in an optical system, may result in large errors in the recorded spectral positions. Any such error can have a significant impact on the accuracy of multivariate statistical classification algorithms that are applied to the recorded spectrum. It is, therefore, important to reduce this error using a calibration procedure that precedes the capture of data.

The sources of error in optical spectroscopy are manifold. The misalignment of the array detector or the diffraction grating can result in error in the wavelength axis. Temperature change in long-term data recording is another key factor in miscalibration [4], which results in expansion of the optical system and minute positional variation of the optical elements. A stepper motor is commonly used in modern spectrometers to rotate the diffraction grating in order to vary the spectral band that is incident on the detector. Such motors provide limited precision in terms of repeatability of the rotation angle, even with closed loop control. Even 0.1° error in the angle of rotation

will result in appreciable error [3]. Another issue that can present when comparing the results from two different spectrometers that have been applied to investigate the same sample, is the individual sensitivity response of each spectrometer. This causes an intensity variation over the bandwidth of the spectrum and can render two spectra of the same sample recorded by two different spectrometers incomparable. Intensity variation across instruments is caused by the differing wavelength and polarisation dependent transmission function of each instrument and possibly differing instrument resolution.

Of particular interest in this thesis is one specific type of spectroscopy known as Raman spectroscopy, which is particularly susceptible to issues related to wavelength and intensity calibration. Raman spectrocopy is based on the inelastic scattering of monochromatic light. [5] The phenomenon was first discovered by Sir C. V. Raman, for which he won the 1930 Nobel Prize in physics. Recent technological developments have made Raman spectroscopy an affordable, non-destructive, and reliable analytical technique. It is possible to identify a specific substance by inspecting the Raman spectrum that is recorded from that substance. For this reason, Raman spectra are sometimes referred to as fingerprints. Raman spectroscopy is utilized in many fields, such as chemistry, physics, biology and medical science. [5–11] Raman spectroscopy has many advantages over other similar methods, including fast detection-speed, repeatability, the requirement for low sample volumes, as well as being non-destructive. Traditionally, Raman systems have been laboratory-based; however, the footprint and cost of optical Raman spectroscopy systems have significantly reduced in recent years; in aviation security for example, spatially offset Raman spectroscopy is now used to rapidly identify material within bottles using a database of spectra recorded from various substances. [6] Raman spectroscopy probes the vibrational and rotational modes of molecules whereby laser photons scattered by the material have lost energy related to the energy of certain Raman-active molecular bonds present within the sample. Raman spectroscopy can identify biomolecular changes within cells as they progress from a healthy to a cancerous state [12–14] making it a powerful technique for the identification of can-

cer cells and tissue. Post-processing such as multivariate statistical analysis [15, 16] is typically applied to Raman spectra for classification, whereby statistical pattern recognition algorithms identify subtle changes across datasets that can be used to accurately differentiate between different pathological groups. [12–14, 17–22]

An application of Raman spectroscopy that is receiving significant attention is the classification of different cell and tissue types, which produce subtly different Raman spectra, due to, for example the presence of disease. [8–11] This involves the training of a statistical algorithm based on known pathological samples. Techniques such as principal component analysis (PCA) and linear discriminant analysis (LDA) are typically employed for the classification of spectra. [8, 9] For example, healthy bladder epithelial cells can be distinguished from low-grade and high-grade bladder cancer cells with greater accuracy using Raman spectroscopy than using traditional techniques for cytology and pathology. [23–25] Similarly, Raman spectroscopy combined with multivariate statistical analysis can be used as a diagnostic tool to detect biochemical changes accompanying cervical cancer [26] as well as oral cancer progression. [27] Raman spectroscopy has the additional advantage of providing rapid minimally invasive detection of the disease, which can be fully exploited using fiber optic probes that facilitate endoscopic classification of tissue, or identification of tumour margins in-vivo. [28, 29] One application of particular commercial interest is Raman guided surgery. [30–32] Another emerging area of clinical research is automated Raman cytology. [33–37]

In order to ensure the accuracy of the collected spectrum, and therefore the accuracy of the resultant classification that is based on this spectrum, for any of the applications that have been discussed above, it is of paramount importance to calibrate the instrument. Unfortunately, there exist many sources of error; a small misalignment in the optical system can result in a significant miscalibration, which can, in turn, lead to incorrect classification of the sample. It is, therefore, necessary to perform an accurate wavelength/wavenumber calibration procedure before starting to record data, and it is often necessary to repeat this calibration step routinely throughout a given experiment; it may be expected that over time, even a well-calibrated system will deviate from its

initial specification; even normal handling can adversely affect calibration. Optical components such as grating, mirrors, lenses and focusing mirror may move slightly over time resulting in miscalibration. In addition, since Raman scattering is a temperature dependent process, it may be expected that variation in ambient temperature will also affect calibration of the recorded spectrum. [5]

Even after wavelength/wavenumber calibration, no two Raman instruments will produce the same raw spectrum for the same sample unless intensity calibration is performed on both systems, and this has hindered progress in many applications including disease diagnostics. Many companies that build Raman spectrometers implement their own internal methods to control the performance and the stability of their own instruments; however, there is no universally accepted method to control the performance and stability of different instruments. The difficulty in comparing spectra that have been recorded across different instruments is one of the main obstacles in the development of Raman spectroscopy for many applications and to its clinical acceptance, and in recent years a number of different protocols have been proposed to address this key issue. [38–43] Intensity calibration and wavenumber calibration are the main subjects of these protocols. This part is concerned only with wavenumber calibration.

Arguably the greatest hindrance to the development of the clinical application of Raman Spectroscopy is the poor cross-instrument comparability, as highlighted by two recent multi-site studies, [44, 45] both of which demonstrate inconsistencies in the wavenumber shift for various materials even following established calibration protocols provided by the instrument manufacturer. Itoh et al. [45] examined spectra from polystyrene, benzonitrile, and cyclohexane obtained across 26 different systems from which they concluded that the wavenumber shift inconsistencies resulted from the instrumentation and calibration protocols and not from the materials samples. Guo et al. [44] found similar inconsistencies across 35 different instruments using acetaminophen, polystyrene and cyclohexane. Cross-instrument differences relate to both wavenumber shift as well as intensity variation for the same sample, the latter being caused by the differing wavelength and polarisation dependent transmission function

of each instrument and possibly differing instrument resolution. Wavenumber variation is generally attributed to small changes in the instrument resulting from thermal expansion or positional drift; instruments with motorised gratings are particularly susceptible to miscalibration. These studies have highlighted the need for further research into the cause of cross-instrument variability and for the availability of open-access standardised materials and calibration protocols that can be universally adopted.

There has already been some development of consensus standards for Raman instrumentation by the American Society for Testing and Materials (ASTM) International in relation to performance testing, calibration, and relative intensity correction (ASTM E1683 [46], E1840 [2], E2529 [47], E2911 [48]). An excellent review of these standards is provided in Refs 49 and 40. In summary, these standards relate to methods for spectral response correction and wavenumber calibration for a single instrument, and for evaluating performance of the instrument in terms of resolution, stray light, sensitivity etc. In the context of this thesis, several of these standards are relevant. For example, Chapter 3, 4, 5, and 7 relate to new methods for wavelength/wavenumber calibration in Raman spectroscopy, and for these chapters, the most relevant of the current set of standards is ASTM-E1840, [2] most recently updated in 2013, which focuses on Raman shift (or wavenumber) calibration. Included in this document are the Raman shift values for eight wavenumber standards including acetaminophen, benzonitrile, which we utilise in Chapters 3 and 5; these values were determined by eight independent laboratories and only the most stable peaks (standard deviation < 1cm^{-1}) were included. Interestingly, this guide does not set out a particular method of calibration; two approaches are commonly used in the literature: (i) wavelength calibration using an atomic spectrum such as from neon followed by wavenumber conversion making use of the laser wavelength, and (ii) direct use of a Raman wavenumber standard such as acetaminophen. In the next chapter, we review both of these approaches in some detail and we conclude that direct wavenumber standards are preferable as they so not require knowledge of the laser wavelength and can provide an accuracy and precision at least as good as wavelength calibration. Chapters 3 and 5 are concerned with novel approaches for direct

wavenumber calibration, and it is our hope that some of our new methods will find their way into ASTM-E1840, when it is next revised. In the context of Chapter 6, the most relevant of the existing standards documents is ASTM-E2911, [48] also most recently updated in 2013, which focuses on Raman intensity calibration using the NIST Standard Reference Materials (SRMs) in the 224X series. However, these standard materials are rarely used in the laboratories due to expense and difficulty in aligning. We hope that the novel method proposed in Chapter 6, which makes use of an arbitrary white light (such as that from the Raman microscope itself) for intensity calibration will also be added to ASTM-E2911 when it is next updated.

The overall goal of this thesis is to develop a range of new methods for wavenumber and intensity calibration, that outperform the existing state-of-the-art, either in terms of accuracy or ease/cost of use, that will push Raman spectroscopy one step closer to clinical adoption. In the next chapter, the background theory relating to Raman spectroscopy and Raman instrumentation is reviewed, which underpins the various contributions in the thesis and detailed literature reviews are provided for both wavelength and wavenumber calibration. Each individual contribution is detailed in a new chapter. Below, these various contributions are briefly previewed, with reference to the chapter in which they are fully documented as well as the associated publications.

1.2 Contributions in this thesis

- **In Chapter 2**, a brief overview of the physics of Raman spectroscopy and Raman instrumentation is provided; this is not considered to be a contribution. However this chapter also contains two valuable literature reviews of wavelength calibration and wavenumber calibration methods, which we believe are contributions to the field. The review of wavelength calibration has been published as a section in the journal paper: *Liu, Dongyue, and Bryan M. Hennelly, "Improved Wavelength Calibration by Modeling the Spectrometer," Applied Spectroscopy (2022): 00037028221111796.* and the review of wavenumber calibration has been included as a section in a recently submitted journal paper: *Liu, Dongyue, and*

Bryan M. Hennelly. "Wavenumber Calibration by Modelling the Raman Spectrometer." submitted to Journal of Raman Spectroscopy Sept 2022.

- **In Chapter 3** we investigate a new Raman wavenumber reference material in the form of a commercial plastic. This material is low cost and highly photostable. It is easy to mount on the microscope and these properties compare favourably with respect to the current standard materials such as benzonitrile or 4-acetamidophenol. This work was published in a conference paper: *Liu, Dongyue, Hugh J. Byrne, Luke O'Neill, and Bryan Hennelly. "Investigation of wavenumber calibration for Raman spectroscopy using a polymer reference." In Optical Sensing and Detection V, vol. 10680, pp. 486-497. SPIE, 2018.* Although, in that paper, we demonstrate that the polymer is not as accurate as 4-acetamidophenol in terms of wavenumber calibration, the accuracy is sufficient for many applications of Raman, and the advantages in terms of cost, photo-stability and ease of use will make this material an attractive option in many cases. There is also the consideration that this material could also be used as part of an intensity calibration routine as described in Chapter 6, and might therefore provide a dual use.
- **In Chapter 4**, we take a step back and focus on wavelength calibration, often a necessary first step in Raman intensity calibration and also for wavenumber calibration (via the wavenumber conversion approach). Here we develop a novel protocol for wavelength conversion, which replaces the less accurate state-of-the-art of polynomial fitting applied to the lines in an atomic emission spectrum such as neon. The approach is to model the relationship between wavelength and the detector pixels using the physics of optical imaging and diffraction. This work was been published as a journal paper: *Liu, Dongyue, and Bryan M. Hennelly, "Improved Wavelength Calibration by Modeling the Spectrometer," Applied Spectroscopy (2022): 00037028221111796.* The advantage of the method over polynomial fitting is to provide greater accuracy, especially in bands where there are few neon lines. We consider the work in Chapter 4 (and the related work in

Chapter 5) to be amongst the most important contributions in the thesis.

- **In Chapter 5**, the work is an extension of Chapter 4. Here, we take the physical model that relates wavelength to detector pixel for an arbitrary spectrometer, and augment it to relate wavenumber to pixel for an arbitrary Raman spectrometer. With this new relationship, the algorithm developed in Chapter 4 can be repurposed to work with a Raman reference spectrum with sharp spectral lines (such as from 4-acetamidophenol of from the polymer in Chapter 3) instead of the neon spectrum. This work has been prepared as a journal paper that has recently been submitted to the Journal of Raman Spectroscopy: *Liu, Dongyue, and Bryan M. Hennelly. "Wavenumber Calibration by Modelling the Raman Spectrometer." submitted to Journal of Raman Spectroscopy Sept 2022.* As in the previous chapter, the advantage of the method over the state-of-the-art is to provide greater accuracy in general, and especially in bands where there are few spectral lines.
- **In Chapter 6**, we change focus and look at intensity calibration. Indeed, this is the only chapter in the thesis to investigate intensity calibration. Typically, intensity calibration is implemented using a reference fluorescent material from NIST (which is rarely used due to expense and difficulties in aligning) or far more commonly using a known NIST-calibrated white light source. The white-light lamp is usually expensive to purchase and needs frequent re-calibration, which is costly and time consuming. In this chapter, we demonstrate how an arbitrary uncalibrated tungsten halogen lamp can be used to achieve higher accuracy than a NIST-calibrated lamp. This work as been prepared as a journal paper to be submitted to the Journal of Raman Spectroscopy with the following title: *Liu, Dongyue, and Bryan M. Hennelly. "Intensity Calibration of Raman Spectrometer using Arbitrary White Light." to be submitted to Journal of Raman Spectroscopy Nov 2022.* There exists a secondary contribution in this chapter, which should not be overlooked; in the course of this work we proposed a novel metric that can be used to quantify the accuracy of an intensity calibration protocol across two or more instruments. Previous attempts to gauge the performance of an intensity

calibration protocol have been qualitative in nature.

- **In Chapter 7**, we return to the subject of wavelength calibration, which was also the topic of Chapter 4. Here we take a fundamentally different approach using a machine learning algorithm known as a long-term short-term memory (LSTM) network. This network has been designed to identify arbitrarily repeating patterns in a training dataset of 1D time-sequences and can then predict the future occurrence of a pattern. This network is the basis of the "Seq2seq" family of machine learning approaches developed by Google for natural language processing. Here we take an off-the-shelf implementation of an LSTM network and apply it to wavelength calibration. The algorithm is trained using hundreds of thousands of neon spectra and their true wavelength axes. A key feature in this chapter is the generation of a simulator that can produce simulated neon spectra using the physical model of a spectrometer developed in Chapter 4. The method is found to be less accurate than traditional third order fitting, except when there are five or fewer neon lines in the spectrum. We believe further improvements can be achieved following this proof-of-concept. This work has been prepared as a conference proceeding in *Liu, Dongyue, and Bryan M. Hennelly. "Wavelength Calibration using Long Short Term Memory Architectures." to be submitted to the proceedings of the SPIE in 2023*.

Chapter 2

Physics of Raman Spectroscopy

2.1 The basic theory for Raman spectroscopy

Raman spectroscopy (RS), named after Indian physicist C. V. Raman (1888 – 1970), is an optical method that probes the molecular structure of a material. Raman spectroscopy is based on the scattering of light caused by molecular vibration and rotation. The wavenumber axis represents the Raman frequency shift, and intensity axis relates to the concentration of the molecule causing the emission. The Raman effect originates from molecular vibration (and lattice vibration) and rotation. When light irradiates a material, both elastic scattering (which is by far the most predominant) and inelastic scattering occur. Elastic scattering has the same wavelength as the excitation light and is called Rayleigh scattering after Lord Rayleigh who first described the phenomenon. Inelastic scattering has components longer and shorter than the wavelength of the excitation light, which are collectively referred to as Raman scattering after C.V. Raman. When photons interact with the molecule, the wavelength of most of the scattered light remains the same, i.e. Rayleigh scattering is abundant. For example, if you point a green laser pointer at a wall, you will always see a green dot; the color of the scattered light is noticeably unchanged. However, although less perceptible, inelastic scattering processes can also occur, resulting in the emission of light of different wavelengths. This is usually related to molecular vibrations and to a lesser extent rotation.

Figure2.1 shows the Jablonski energy level diagram which relates to quantum me-

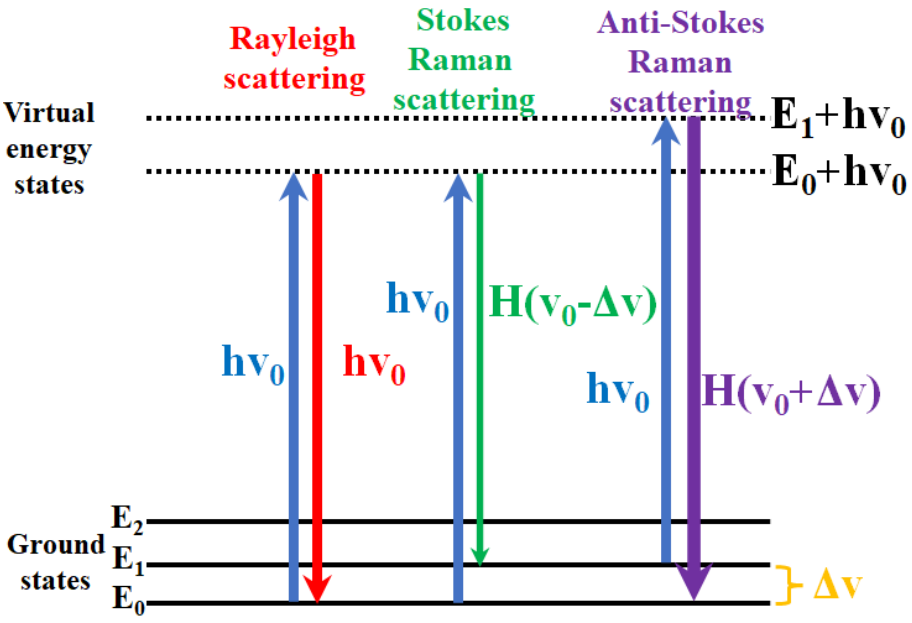


Figure 2.1: Jablonski energy level diagram for Rayleigh and Raman scattering.A Raman scattering event (or resonance) produces a Lorentzian line shape in the Raman spectrum.

chanical energy states. By comparing the energy differences of the photons emitted during Rayleigh scattering and Raman scattering, it is shown that Raman scattered photons either lose energy (Stokes) or gain energy (Anti-Stokes) during this scattering process. The molecule moves from its initial state E_0 or E_1 up to a virtual state $E_0 + h\upsilon_0$ or $E_1 + h\upsilon_0$ due to the energy gained from the incident photon. If the state from virtual states then drops back to the initial state, the collision is elastic and the released photon and energy/wavelength are identical to that of the laser. However, in some rarer cases, the drop will be to a different vibrational/rotational state resulting in the emission of a photon with a different energy/wavelength. In those cases where the new state has lesser energy than the initial state, Stokes Raman photons are emitted at $H(\upsilon_0 - \Delta\upsilon)$, and in the cases where the new state has higher energy, anti-Stokes Raman photons are emitted at $H(\upsilon_0 + \Delta\upsilon)$. The ratio between the Stokes and anti-Stokes emissions relates to the temperature, and typically at room temperature, the Stokes emission is measured by a Raman spectrometer.

The physics of Raman scattering should also consider the quantum effect of the molecule. The molecule has vibrational and rotational degrees of freedom under the

quantum effect. Even if the molecule stays in a static electric field and is not illuminated by electromagnetic radiation, the ground state still exists in lots of different energy levels. Figure 2.1 oonly shows three energy levels in the ground state for brief expression, but it does not mean there are only two levels. In fact, lots of energy quantum effects are created by the molecule's vibration and rotation in the ground state. Usually, the final state of the Rayleigh scattering has the same electronic energy as the initial state, and the vibrational energy in Stokes Raman scattering has a higher final state, and lower in anti-Stokes Raman scattering. Based on the Boltzmann statistics, [50] the intensity ratio between the Raman Anti-Stokes and Raman Stokes $\frac{\upsilon+\upsilon_R}{\upsilon-\upsilon_R}$ in the settled temperature T can be present as:

$$\Delta\tilde{\nu}(cm^{-1}) = \left[\frac{1}{\lambda_{incident}(nm)} - \frac{1}{\lambda_{scattered}(nm)}\right] \times \frac{10^9(nm)}{10^2(cm)} \tag{2.1}$$

The first term is the wavenumber Raman shift in cm^{-1}, $\lambda_{incident}$ is the wavelength of the excitation laser in nm, and $\lambda_{scattered}$ is the wavelength of the Raman scatter in nm. The magnitude of the emission is dependent on the rotational and vibrational energies of the molecules, as will be discussed in more detail in Sections 2.1.1 and 2.1.2.

2.1.1 Rotational Raman spectroscopy

The molecule must be anisotropically polarizable such as for the case of O_2, N_2, and H_2; molecules that exist in permanent dipole moments will not undergo rotational Raman scattering. [51] The term "anisotropically polarizable" refers to a property of a molecule or material that can be polarized to different extents in different directions, resulting in anisotropic behavior. In other words, the polarizability of the molecule or material depends on the direction of the electric field to which it is exposed. This property is particularly relevant in the study of light scattering phenomena, where the polarizability of the material determines the strength and direction of the scattered light. Anisotropic polarizability can be described mathematically using a tensor, which relates the induced dipole moment of the molecule or material to the applied electric field. The polarizability of the molecule will determine the impact of an externally applied

electromagnetic field on the electron charge distribution of the molecules. An electric field such as that produced by laser radiation will induce a dipole moment in molecules that are polarizable. In such an electrostatic field, inside the molecule the positive nucleus will move towards the negative pole of the field, and negative electrons move towards the positive pole of the field. In general, the induced dipole moment of the molecule is proportional to the effective electric field strength E acting on it, that is:

$$\mu_{induced} = \alpha E \tag{2.2}$$

where $\mu_{induced}$ is the size of induced dipole moment, E is the amplitude of the electric field, and α denotes polarizability of the molecule. The polarizability of the molecule α can be described as the degree to which the electrons in the molecule can be distorted in a static electric field.

We now consider the molecule in the electromagnetic field from a monochromatic incident laser in more detail; the electric field can be described as a time dependent function, $E(t)$:

$$E(t) = E_0 \cos(2\pi \upsilon t) \tag{2.3}$$

where E_0 is the amplitude of the electric field from the laser, υ is laser frequency, and t is time. Thus, the induced dipole moment(μ), in Equation 2.2, can be rewritten as:

$$\mu = \alpha E(t) = \alpha E_0 \cos(2\pi \upsilon t) \tag{2.4}$$

Most of emitted photons will have the same oscillation frequency υ, as the monochromatic incident beam, which can be explained by Rayleigh scattering. Molecular rotation can also affect some of the emitted photons. The polarizability of the molecule can be related to molecular rotations, and the electric field can therefore produce rotational oscillation. If a molecule is rotationally oscillating at a frequency of υ_R, variations in the polarizability ($\Delta\alpha$) will be time dependent:

$$\alpha(t) = \alpha_0 + \Delta\alpha \cos(2\pi \upsilon_R t) \tag{2.5}$$

where α_0 is a constant representing the average polarizability in one cycle of a rotational oscillation; $\Delta\alpha$ represents the polarizability, which is time-varying in one cycle of molecular rotation. With α ranging from $\alpha_0 - \Delta\alpha$ to $\alpha_0 + \Delta\alpha$ as the molecule rotates through 2π. By substituting the expression for polarizability (Equation 2.5) into the expression for induced dipole moments (Equation 2.4), the induced dipole moment can be expressed:

$$\begin{aligned} \mu = \alpha E(t) &= [\alpha_0 + \Delta\alpha \cos(2\pi \upsilon_R t)] \times E_0 \cos(2\pi \upsilon t) \\ &= \alpha_0 E_0 \cos(2\pi \upsilon t) + E_0 \cos(2\pi \upsilon t) \Delta\alpha \cos(2\pi \upsilon_R t) \end{aligned} \tag{2.6}$$

$$\mu = \alpha E(t) = \alpha_0 E_0 \cos(2\pi \upsilon t) + \frac{1}{2}\Delta\alpha E_0 [\cos 2\pi(\upsilon - \upsilon_R)t - \cos 2\pi(\upsilon + \upsilon_R)t] \tag{2.7}$$

where the first part $\alpha_0 E_0 \cos(2\pi \upsilon t)$ represents the Rayleigh scattering, the latter two terms oscillate at the sum and difference frequencies, representing the Stokes Raman scattering and anti-Stokes Raman scattering, respectively. If $\Delta\alpha = 0$, the molecule is not polarizable, which means the anti-Stokes and Stokes components disappear. In order for rotational Raman scattering to occur a change in the polarizability caused by molecular rotation is required.

The physics of Raman scattering can also be interpreted in terms of quantum mechanics. The molecule will have vibrational and rotational degrees of freedom under the quantum effect. Even if the molecule stays in a static electric field and is not illuminated by the electromagnetic radiation, many different energy levels will exist at the ground state. In Fig. 2.1 only three levels are shown in the ground state for simplicity, but in general there will exist many more. Based on the Boltzmann statistics, [50] the intensity ratio between the Raman Anti-Stokes and Raman Stokes $\frac{\upsilon+\upsilon_R}{\upsilon-\upsilon_R}$ in the settled temperature T can be present as:

$$\frac{\upsilon + \upsilon_R}{\upsilon - \upsilon_R} = \exp\left(\frac{-h\upsilon_R}{k_B T}\right) \tag{2.8}$$

where υ_R is frequency of Raman shift, h is the Planck's constant, k_B is the Boltzmann's constant and T is the temperature associated with the scattering species. This equation is sometimes used to measure the temperature via Raman spectroscopy. [52] It can

be shown using Equation 2.8, that the Stokes Raman scattering is stronger than the anti-Stokes Raman scattering at room temperature. [53]

2.1.2 Vibrational Raman spectroscopy

The classical description of vibrational Raman spectroscopy is similar to that presented for rotational Raman spectroscopy; here, molecular vibration can also lead to polarizability changes. [51] For polyatomic molecules, vibrational activity is complex, often requiring the application of group theory to determine if vibrational modes exist and, therefore, whether a molecule is Raman active or not as described in Reference 54. In general, two rules can be applied to determine whether a polyatomic molecule undergoes Raman vibrations: (i) if the molecule has no symmetry (e.g. HCl, HCN), then in general all of its vibrational modes are Raman active; however, these asymmetric vibrations are weak and usually not observable; (ii) if the polyatomic molecules are symmetric (such as H_2O, CO_2), then the corresponding vibrations will produce strong Raman scattering.

A vibrational Raman spectrum contains the unique and highly resolved vibrational signature of the scattering that is typically associated with a Raman spectrum. In the measurements, the Stokes part is commonly much more intense than the anti-Stokes part in an entire spectrum, [55] as described earlier, and the spectrum is typically recorded in the Stokes band. The vibrational Raman scattering is similar to the rotational Raman scattering mentioned in Section 2.1.1; however, during the period of the molecular bond vibration, the electron distribution is stretched and compressed periodically, which gives an oscillation of the polarizability along the direction of the electric field. Based on Equation 2.5, it can be shown that the induced dipole oscillates in phase with the vibrational motion of the molecule:

$$\alpha(t) = \alpha_0 + \Delta\alpha \cos(2\pi \nu_{vib} t) \tag{2.9}$$

where $\Delta\alpha$ is the amplitude of the change in the polarizability varying by vibrational cycle, ν_{vib} is the frequency of the vibrational oscillation, α_0 is a constant that represents

the average polarizability in one cycle of vibrational oscillation, and t is time. Similar to Equation 2.7 for rotational Raman spectroscopy part, a time dependent induced dipole moment is described as follows:

$$\mu = \alpha E(t) = \alpha_0 E_0 \cos(2\pi \upsilon t) + \frac{1}{2}\Delta\alpha E_0 \left[\cos 2\pi(\upsilon - \upsilon_{vib})t - \cos 2\pi(\upsilon + \upsilon_{vib})t\right] \quad (2.10)$$

where $\upsilon + \upsilon_{vib}$ indicates the frequency of the emitted anti-Stokes Raman scattering, and $\upsilon - \upsilon_{vib}$ is the frequency of emitted Stokes Raman scattering. As for the case of rotational Raman spectroscopy, Rayleigh scattering, $\alpha_0 E_0 \cos(2\pi \upsilon t)$, dominates and has the same frequency as the incident light.

From the description in Section 2.1.1 and Section 2.1.2, it is clear that Raman scattered photons can be generated in two different ways: by rotation scattering and/or vibrational scattering. The scattered photons generated by these scattering events are collected by a spectrometer and manifest as intense peaks at different wavelengths. Wavenumber shift is the unit of position of the Raman spectrum and can be related to wavelength relative to the wavelength of the laser according to Equation 2.1.

2.1.3 Some limitations of Raman spectroscopy

Raman spectroscopy has several advantages over other forms of spectroscopy. Firstly, it is label free and can be applied without damaging the sample. It can be applied non-invasively to probe the chemistry of living samples as well as testing the purity of chemicals. It can be applied with μm resolution using a microscope objective to deliver/focus the laser and to collect the scattered photons; it can also be used endoscopically using optical fibers to do the same. It can provide fingerprint like identification of pharmaceutical purity and diseased cells/tissue. However, Raman spectroscopy also has some well known limitations including the following:

1. It cannot in general be used for metals or alloys.
2. The Raman effect is very weak, and detection requires sensitive and highly optimized instruments. This also necessitates long acquisition times when capturing a Raman spectrum from a biological specimen and prohibits the capture of hy-

perspectral images over large areas.

3. Increasing the power of the laser will increase the amount of Raman scattering proportionally and this approach can be used to reduce the acquisition time but only up to the point where the sample is damaged by heat.
4. The overlapping of different vibrational peaks and the intensity of Raman scattering are easily affected by the resolution of the spectrometer, which can complicate cross-instrument comparability.
5. The fluorescence from impurities in glass and similar materials can result in strong signals that can dominate the Raman spectrum. Glass lenses can cause problems especially for excitation in the NIR. The use of a confocal aperture to spatially resolve the source of the scattering to be matched with the focal point of the microscope objective (in a Raman microscope) can help to reduce the presence of background signals.
6. Similar to the previous point, the sample container or substrate (eg glass slide) can be especially problematic and since it is in close proximity to the sample, the use of a confocal aperture cannot solve this problem entirely. Various algorithms have been developed to numerically subtract the glass fluorescence spectrum but when NIR excitation is used, the Raman spectrum is relatively weak compared to this signal. This is less of a problem for visible wavelengths. Another solution is to use substrates that produce a weak background; however, such pure Calcium Fluoride or quartz crystals are expensive and cannot be used as consumable for clinical applications. An example is shown in Section 5.5.1 in which the glass cuvette sample holder is drilled at the base and sealed with a Calcium Fluoride coverslip to be used with an inverted Raman microscope. This coverslip produces a negligible Raman/fluorescence spectrum and the chemical with the holder can be probed in isolation.
7. The highly variable auto-fluorescence from a biological sample can also be problematic as it adds a variable baseline to the spectrum and makes difficult cross-instrument comparability. This effect is more pronounced for visible wavelength

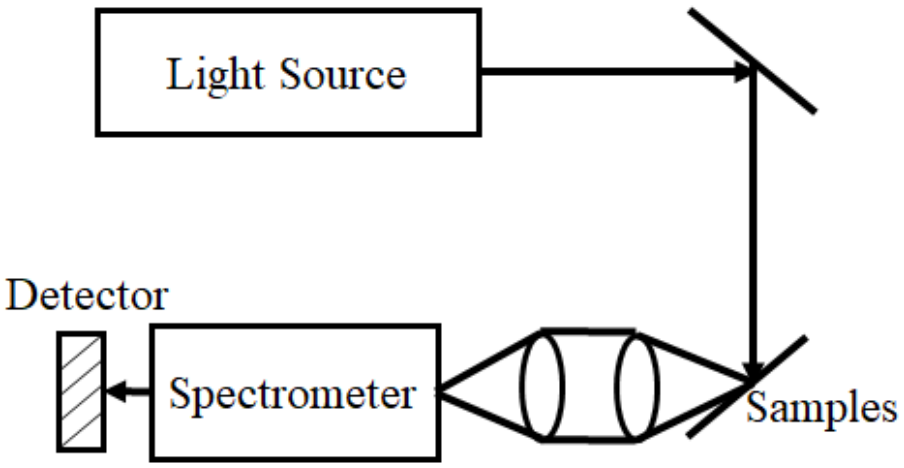


Figure 2.2: A basic Raman Spectrometer

excitation compared with NIR. As for removing the glass spectrum, various methods have been developed to numerically subtract this baseline, see Chapter 6 on intensity calibration.

8. Classification/identification of spectra recorded across different instruments is only possible if careful intensity and wavenumber calibration is performed. This is the subject of this thesis.

2.2 Optical system of Raman spectroscopy

The probability of Raman scattering is extremely small relative to Rayleigh scattering, and the strongest Raman scattering only accounts for a few thousandths of the entire scattered light. Since the Raman effect itself is relatively weak, the optical components of the Raman spectrometer must be optimized to deliver the laser to the sample, and the scattered light to the spectrometers with the highest efficiency possible.

Figure 2.2 shows an example of a basic Raman spectroscopy setup, which consists of four main components: (i) an excitation source (i.e. laser or filtered emission lamp), (ii) an optical system to deliver the light to the sample with minimum power loss, (iii) a collection system with highly efficient optics to collect an optimum amount of Raman scattered light, and (iv) a detection system (i.e. spectrograph and detector). Optical system design is an essential part of a successful Raman experiment. The laser source must be efficiently delivered to the sample while simultaneously, an efficient optical

system must be designed to collect as many Raman photons as possible. To achieve this, a series of mirrors and lenses are used in the light path for both delivery and collection, and filters are used to remove unwanted signals. Filters are essential in the collection path to reduce the dominant Rayleigh scattering signal and other noise sources. Finally, the spectrum itself is dispersed via a grating and is recorded by a cooled CCD. Noise in the CCD device is caused by thermal energy. In order to get high signal-to-noise ratio (SNR), the CCD must be cooled. The slit width at the entrance to the spectrometer is an additional consideration; a wider width will increase power throughput (and lower acquisition time) but at the expense of negatively impacting spectral resolution. More details about the Raman spectrometer designed used in this paper will be shown in Section 5.2.2. In the next few subsections, I will focus on several important components of the Raman optical system, including the light source, spectrometer and CCD camera, and explain how these components affect the performance of Raman spectroscopy.

2.2.1 Light source

There are many types of light sources used in Raman systems. Incoherent light sources include filtered incandescent lamps (which were used in the earliest experiments) and light emitting diodes (LEDs), and coherent light sources include various lasers. Lasers can be divided into gas lasers, liquid lasers, solid-state lasers and semiconductor lasers. Solid state lasers have become the source of choice in modern Raman spectroscopy due to their decreasing cost and narrow linewidth. In general, Raman spectroscopy is independent of the wavelength of the laser; however, the choice of wavelength can have an impact depending on the sample of interest. If you study biological proteins, cells, etc., you need excitation using longer wavelengths of near-infrared light to avoid auto-fluorescence from the sample; however this prohibits the use of glass substrates or sample containers. Visible wavelengths and glass substrates can be used for biological samples if the auto-fluorescence and glass signals can be numerically subtracted by some method (see examples in Chapter 6). In this thesis, two excitation laser wavelengths were used a 785nm CLDS diode laser with power of 300mW and a 532 nm solid state diode laser with a power of 50 mW.

An additional consideration with respect to wavelength is the number of Raman scattered photons. The intensity of the collected Raman photons I is proportional to the fourth power of the incident laser frequency v in the following equation:

$$I \propto v^4 \tag{2.11}$$

The Raman intensity scales linearly with respect to laser power. According to Equation 2.11, we can determine that the intensity ratio of the collected Raman photons generated by a 532 nm laser and an 830 nm laser can reach to 5.92. Therefore, for a camera with uniform quantum efficiently, the acquisiton time using 830 nm excitation will be six times longer to obtain the same spectrum recorded with 532 nm.

Excessive thermal energy from the laser can damage biological and chemical samples. It is important to note that both flammability and toxicity are serious considerations when selecting appropriate chemical samples for wavenumber calibration (see Chapter 3 and Chapter 5). The high temperature generated by the laser can cause chemical samples or volatile toxic gases to be ignited. Generally such samples are avoided, and this is certainly true in this thesis. For biological samples, the typical thermal damage is coagulation, which is manifested by the irreversible transformation of liquids within cells and proteins into a gel state [56] if the temperature caused by the laser exceeds approximately 100°C.

2.2.2 Spectrometer and CCD camera

In a Raman spectroscopy setup, a spectrograph is employed which combines a spectrometer and a CCD camera, which is used to record a Raman spectrum. In a typical Raman spectrometer, Raman scattered light is dispersed using a diffraction grating. This scattered light is projected onto the long axis of the CCD array. In Figure 2.3(b), a reflection spectrometer is shown. The Raman scattered light is first focused into the entrance slit through the lens. After passing through the slit, the scattered light is transformed into parallel light through the collimating concave mirror and hits the diffraction grating surface; reflective diffraction gratings are reflective plates with periodic arrays of

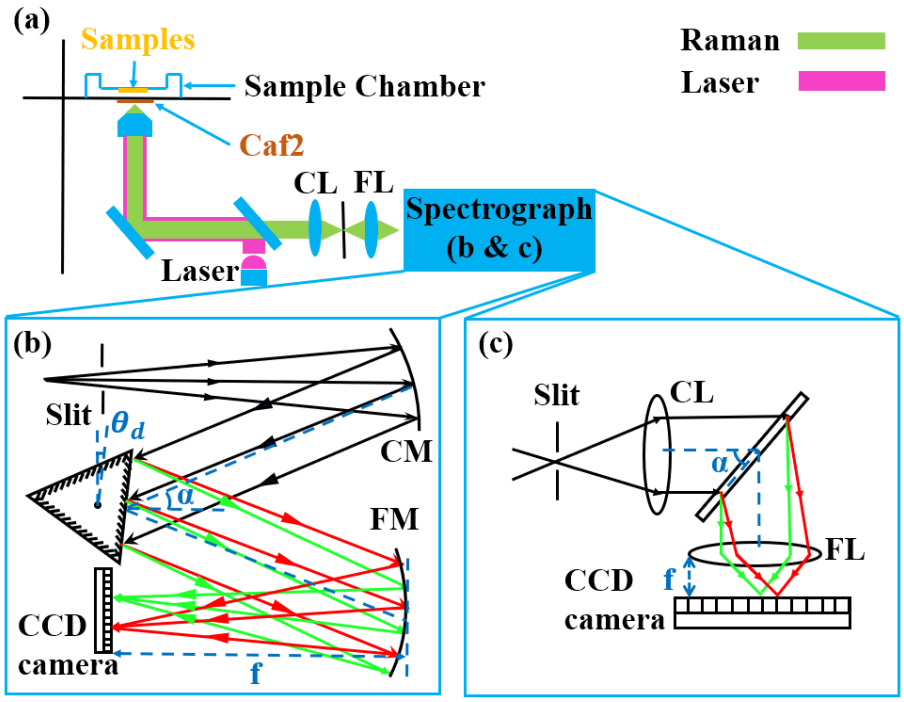


Figure 2.3: (a) Basic Raman spectrometer with epi-illumination; (b) A Czerny-Turner spectrograph with a rotating grating; the parameters shown in the illustration appear in Equation 5.1 in the text; (c) A transmission spectrograph using a holographic grating. Both types of spectrographs are used in this study.CM: Collimating mirror; FM: Focusing Mirror; CL: Collimating Lens; FL: Focusing Lens

grooves on the surface. The concave mirror reflects the light separated from the grating to different pixel positions of the CCD camera. The CCD detector contains an array of light sensors capable of detecting light with spatial resolution. It is divided into a large number of small photosensitive regions called pixels. Because they are extremely sensitive to light, this makes these detectors suitable for analyzing inherently weak Raman signals. Each pixel in a CCD camera is an individual photodiode, which converts the collected photons into electrons and outputs a digital signal to represent the collected intensity. The CCD must be cooled before experiment starts to ensure that thermal noise generated by thermal energy by the pixels can be suppressed. Figure 2.3(c) shows the transmission spectrometer system, which employs a holographic diffraction grating in a transmission architecture to disperse the light. In this cases glass lenses are used to collimate the light from the slit onto the grating and also to focus the dispersed light to the camera. Both types of spectrometer are employed in Chapters 4 and 5 and further details are presented in these chapters on modelling these spectrometers.

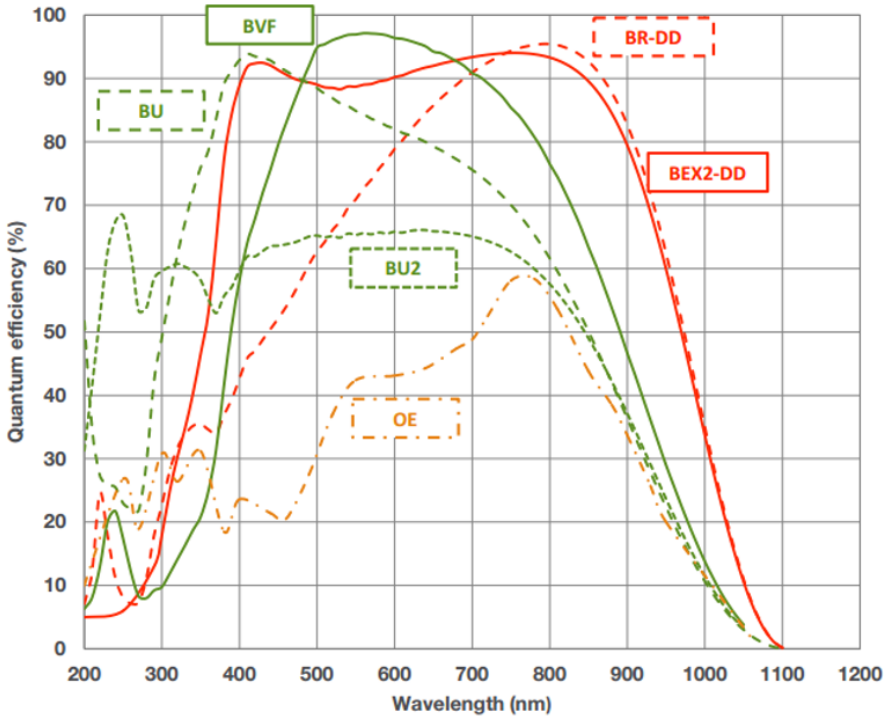


Figure 2.4: QE of BR-DD CCD (red-dotted line). This is the QE of the camera that is used in the Raman micro-spectrometer in MU biophotonics lab.The QE of the camera modulates the Raman intensity and therefore two cameras can produce appreciably different Raman intensities.

The Quantum Efficiency (QE) of the camera is related to the ability of the sensor to respond to the incoming photon signal wavelength and the conversion of it to a measurable electron signal, which is shown in Figure. 2.4. It also affects the accuracy of intensity of the spectrum across the recorded bandwidth, as it highly wavelength dependent for a given camera.

2.3 Cause of miscalibration in terms of wavelength and wavenumber

There are many sources of error that can lead to miscalibration of a Raman spectrum. Slight rotation of the spectrograph diffraction grating or the CCD (for the case of the CCD we mean in-plane rotation), small lateral displacement of the CCD and small changes in the laser excitation wavelength due to variation in temperature or current will all result in errors in terms of wavenumber position of spectral peaks. The most

potent source of error in many Raman systems is the grating angle. Many spectrographs allow for rotation of the grating in order to allow for recording different regions of the spectrum. However, frequent rotation of the motor can lead to error. The best way to correct for this error is to routinely perform a rigorous calibration procedure. In this section, we examine these various sources of error and attempt to relate instrument error to error in the wavenumber axis in the recorded spectrum.

A common feature in a modern spectrograph is the variable rotation of the grating in order to vary the wavenumber band that is recorded on the CCD camera; this is usually achieved using a stepper motor. However, no motor can provide precise repeatability in terms of the rotation angle. As a result, it is difficult to ensure that the expected angle is obtained. Figure 2.5 illustrates the effect of an error in the angle of 0.1 degrees in terms of the resulting error in both wavelength, see Fig.2.5 (a), and in wavenumber, which is shown in Fig.2.5 (b). This calculation is based on the grating equation and the optical configuration of a Czerny-Turner spectrograph such as that used in the experimental system described in Section 4.3. The parameters used in this calculation are similar to those found in the experimental system: source wavelength is 532nm, spectrograph focal length is 0.8m, CCD pixel size is 26 μm, and grating incidence angle is $\theta i = 11°$ and groove density of the grating is 600 lines/mm. Fig.2.5 (a) shows the error in wavelength that occurs when an expected angle of 11° is used to calculate the wavelength axis for the spectrum, and the actual grating angle is 11.1°; the resulting error in the wavelength axis is approximately 2.83 nm for all pixels. In Fig.2.5 (b) the corresponding error is shown for the wavenumber axis; this error varies from 100 cm^{-1} to 80 cm^{-1} across the spectrum.

The second source of error that is considered here is unexpected displacement of the camera, which may occur if the camera is replaced. The error in both the wavelength and wavenumber positions of spectral peaks resulting from a shift of 0.26 mm (10 pixels) is shown in Fig.2.5 (c) and Fig.2.5 (d), respectively for the same parameters used in the previous example. This number(10 pixels) is selected arbitrarily. In Fig.2.5 (c) it can clearly be seen that shift 10 camera pixels will result in approximately 0.5nm error

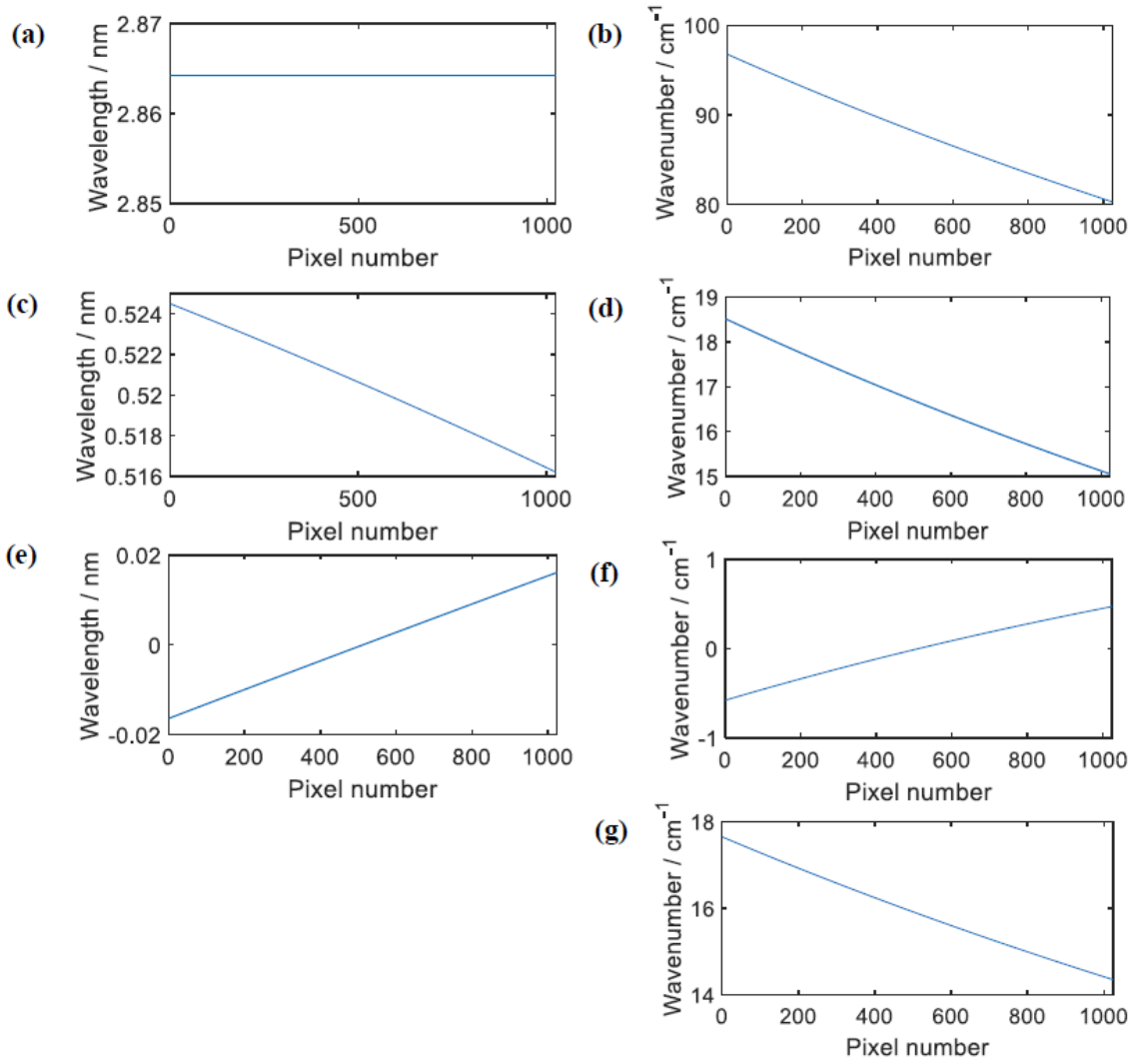


Figure 2.5: (a) shows the wavelength error caused by the 0.1-degree rotation in the grating as a function of pixel number and (b) shows the corresponding error in wavenumber; (c) shows the error in the wavelength axis as a function of pixel number, caused by a lateral displacement of the camera by an amount equal to the width of 10 pixels (approximately 0.25mm) and (d) shows the corresponding error in wavenumber; (e) shows the error in wavelength as function of pixel number caused by an in-plane rotation of the camera sensor by an angle of 2° and (f) shows the corresponding error in wavenumber. Finally (g) shows the error in wavenumber caused by a shift in the source laser wavelength from an expected value of 532 nm to a value of 532.1nm.

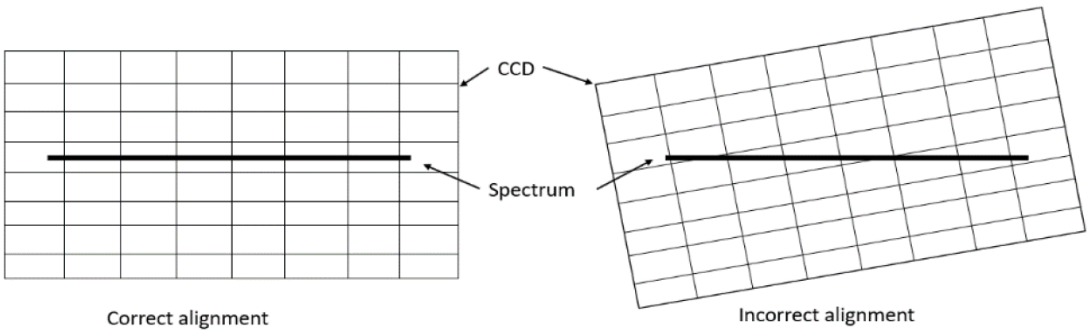


Figure 2.6: Illustration of the error caused by incorrect placement of the CCD camera in terms of in-plane rotation. The angle of rotation shown in the figure is exaggerated.

for every point in the spectrum; more accurately, the error varies from 0.515nm to 0.525nm across the spectrum. Similarly, in Fig.2.5 (d) it can be seen that in terms of wavenumber, this corresponds to an error of between 15cm^{-1} and 18.5cm^{-1} across the face of the camera. It can be concluded that a small shift of the camera position results in a significant miscalibration of the wavenumber position of spectral peaks and may lead to significant errors in terms of multivariate classification. Another camera related source of error that may occur is a small unexpected in-plane rotation, which may occur due to the slight incorrect placement of the camera in the output port of the spectrograph as illustrated in Figure 2.6. This will reduce the effective width of the CCD pixels relative to the plane of the diffraction grating.

If the CCD camera is rotated around its center, the wavelength/wavenumber error at the center pixel number should be zero and will increase at increasing distances from this point. In Fig.2.5 (e) the error in wavelength is shown as a function of pixel number for an incorrect placement angle of 2°. The error is zero for the centre pixel but increases to ±0.02nm at the ends of the camera. The corresponding error in wavenumber is shown in Fig.2.5 (f); once again, the error is zero for the centre pixel but increases to ±0.5cm^{-1} at extreme ends of the camera. Although this is not as potent a source of error as grating rotation or lateral camera displacement, it is, nevertheless, noticeable.

When calculating the wavenumber axis using an already available wavelength axis, the laser wavelength is a key factor. A slight error of laser wavelength can lead to the large difference in the resulting wavenumber axis. Modern solid stable lasers are controlled using highly stabilized laser drivers, which guarantee that a single longitudinal mode is

produced; however, the wavelength may vary slightly over long periods of time. In Fig.2.5 (g) the resulting error in wavenumber is shown when an expected laser wavelength of 532nm was used to calculate the wavenumber axis, but the actual laser wavelength is 532.1nm; the resulting error varies from 14cm^{-1} to 18cm^{-1} across the spectrum. Even a small change in wavelength of 0.001nm during an experiment will lead to appreciable changes in the wavenumber axis; to overcome this problem, many of the lasers used for Raman spectroscopy have picometer wavelength stability over a number of hours.

The process of Raman scattering is dependent on temperature. [4] Significant changes in temperature can lead to fundamental changes in molecular structure. [4] A change in temperature can, therefore, change a Raman spectrum in ways that are sample dependent; peak broadening can occur, as can changes in peak intensity; however, such changes are not typically reported for the types of samples used in the applications discussed in Section 3.3, over the range of normal ambient temperatures that can be expected in a laboratory environment. However, changing temperature may also result in miscalibration of the wavenumber axis by a few wavenumbers, which may be due to slight thermal expansion of the optical elements. Sample heating during Raman spectroscopy can have several impacts on the measurement and interpretation of Raman spectra:

1. Changes in the sample composition: High temperatures can cause chemical changes or decomposition of the sample, leading to changes in its Raman spectra. This can lead to incorrect interpretation of the spectra or loss of important information.
2. Signal intensity changes: Heating the sample can increase the intensity of Raman signals, making it difficult to distinguish the real signals from noise. This can result in incorrect peak assignments or overestimation of the concentration of certain components.
3. Peak shifts: Heating the sample can cause peak shifts due to thermal expansion or contraction of the sample or changes in its refractive index. This can lead to peak overlapping and difficulties in peak identification.

4. Laser-induced heating: The laser used in Raman spectroscopy can also cause heating of the sample. This can lead to thermal changes in the sample and affect its Raman spectra.

Therefore, to avoid the impact of sample heating during Raman spectroscopy, it is important to keep the sample temperature stable and low, and avoid using laser power that is too high. It is also important to understand the thermal behavior of the sample and to carefully interpret the Raman spectra obtained under different temperatures. The problem can be mitigated by controlling room temperature. It is not straightforward to simulate an error in wavenumber due to temperature change; instead, an experiment was designed to investigate this effect. A polymer slide, discussed in more detail in Section 3.3, was illuminated by a 532nm wavelength laser source using the experimental set-up described below, with no temperature control. The position of one peak was measured for 1000 consecutive 10s acquisitions. The wavenumber position of this peak as a function of the sequence number is shown in Fig. 2.7; an increase in the mean wavenumber position of almost 2cm^{-1} is measured from the beginning of the experiment to the end; ambient temperature was recorded to have increased by approximately 3°C during this time. We note that this effect may be the result of other sources of error; a more controlled experiment would be required to conclusively relate temperature change to instrument miscalibration.

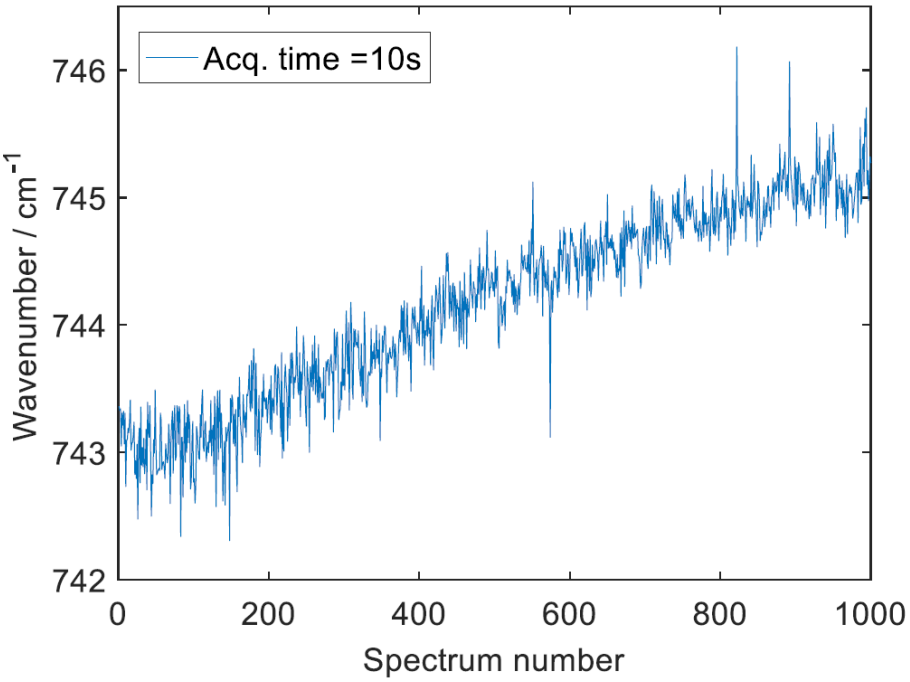


Figure 2.7: Change in wavenumber position at spectral peak for a polymer sample over time.

2.4 Wavelength calibration: Background

We note that the review that follows is included as part of a paper, *Liu, Dongyue, and Bryan M. Hennelly. "Improved Wavelength Calibration by Modeling the Spectrometer." Applied Spectroscopy (2022): 00037028221111796.* recently published.

2.4.1 Wavelength calibration using polynomial fitting

Wavelength calibration of a spectrometer using a detector array is based on exploiting the relationship between wavelength and pixel position across the detector using wavelength reference standards, such as neon or krypton, which have well defined peak wavelengths. [57, 58] Typically, this involves fitting a low-order polynomial to pixel-position and wavelength coordinates for a series of known peaks in the reference. The use of linear and higher order polynomials have previously been applied for this purpose; the selection of polynomial order varies in the literature on a case by case basis. Here, we provide a brief review of the key contributions in this area in recent decades, and in later sections the contribution proposed in this chapter is described in the context of this background material.

In the late 1980s, Hamaguchi proposed a method for the calibration of Raman spectrometers. [59–61] At that time, the use of 'multi-channel detectors' was relatively new and included instruments such as silicon-intensified target tubes, intensified photodiode arrays (IPDA) and early-stage charge coupled devices (CCD) with limited extent. The basis of Hamaguchi's approach was to first perform wavelength calibration using a wavelength standard such as neon, followed by conversion to wavenumber, making use of the laser wavelength in this calculation. In the simplest case, in the absence of distortion, a linear relationship between 'pixel' and wavelength was assumed and a least-squares approach was proposed in order to achieve accurate calibration using only a few neon peaks. However, it was also emphasised in this work that 'optical-distortion' caused by spherical aberration in the spectrometer, or by the detector itself, such as in the case of an electrostatic-type IPDA, could result in a pincushion effect and a non-linear relationship between wavelength and position in the recorded spectrum.

Hamaguchi proposed a solution to this problem, whereby a wavelength standard with many peaks (such as a neon lamp in an appropriate band) could be recorded and a higher order polynomial could be used to describe the relationship between the recorded peak (distorted) positions, and the expected peak positions. Thereafter, recorded spectra would be first corrected for the non-linearity caused by the optical distortion by using this predetermined higher order polynomial to cast the spectrum into a 'virtual channel'. Following this, a linear-relationship between wavelength and position could be assumed, facilitating a least-squares fitting of straight-line approach to calibration as in the simple case in which no distortion was present. Importantly, Hamaguchi notes that in the case of a non-linear relationship between wavelength and position, a large number of reference peaks are required and, furthermore, the reference spectrum should contain peaks close to each end of the spectrum, since a least squares approach with higher order terms will often provide erroneous results when the fitted curves are extrapolated to regions where no data points are available.

In this chapter, we also propose to account for the non-linearity of wavelength and pixel positions using a non-linear relationship; however, we do not limit ourselves to the use of fixed order polynomials. Instead, we model the spectrometer using basic diffraction theory and ray optics in order to derive the non-linear relationship. Like the Hamaguchi method, we cast the recorded wavelength-pixel positions of several neon peaks using this non-linear relationship, such that the relationship between wavelength and position becomes linear, followed by least-squares fitting of a straight-line. This method accounts for non-linear dispersion by the grating but does not attempt to account for optical-distortion as for the case of the method described above.

Linear/first order fitting has also been applied to splice together adjacent spectral bands, [62, 63] and has also been applied as the first-step in intensity calibration using a calibrated white lamp or florescence standard, which is used to correct for variation in spectral intensity caused by wavelength variable transmission of the optical elements in the spectrometer or the wavelength dependent efficiency of the grating. [63] This method assumes a linear relationship between position and wavelength, which is ap-

proximately true over narrow spectral bands and for low dispersion gratings. Calibration using linear regression is known to produce errors as a consequence of the non-linear relationship between wavelength and pixel position, which becomes more pronounced for high dispersion gratings. [64–67] For some applications, such as for splicing, and for intensity calibration, or indeed for calibrating low dispersion systems, these errors are small enough to have low impact. However, for more accurate characterisation of wavelength positions, up to fifth-order fitting has been preferred in some cases. [62, 68]

Tseng et al. established possibly the most widely adopted protocol for wavelength calibration of modern spectrometers. [62] Included in this protocol is the use of first-order fitting as a means to stitch together adjacent spectral windows, as well as second-order fitting in order to obtain higher accuracy. This protocol also included a method to improve results by first interpolating the peak regions in the spectrum in order to obtain sub-pixel accuracy of peak position. The authors reported a standard-deviation in the calibrated wavelength positions of the neon peaks < 0.005nm for a 1800 lines/mm grating and a spectrometer with 0.64m focal length.

Despite the better accuracy provided by second-order fitting, some groups have continued to use first-order fitting of wavelength and pixel position. Hutsebaut et al. have established a widely adopted protocol for the calibration of a Raman spectrometer. [39] For intensity calibration, they record a neon wavelength standard followed by first-order fitting of the peak wavelengths and pixel positions. This is used as a first step in order to wavelength-calibrate a white light reference spectrum, which is subsequently used for the intensity calibration of a Raman spectrum recorded using the same spectrometer. Since the intensity of this reference is relatively smooth with respect to wavelength, the accuracy afforded by linear-fitting is sufficient. As an indicative value for the goodness of fit, the Root-Mean-Square-Error (RMSE) was calculated by the authors to be 0.03nm for the calibrated neon wavelength values.

Carter et al. proposed three methods of Raman wavenumber calibration, [69] one of which is based on wavelength calibration using a neon reference with first-order fitting of peak wavelengths and pixel position, followed by wavenumber conversion using

the known wavelength of the excitation laser. The authors argue that their approach is simpler to the protocol in Ref. 62 and is, therefore, more suitable for frequent re-calibration. First-order fitting of wavelength and pixel is shown to be sufficient for the calibration of relatively narrow Raman bands (100cm^{-1}) and the authors state that higher-order fitting would be preferable for wider bands as outlined in Ref. 62.

Gaigalas et al. employed first-order fitting of wavelength and pixel position as a first step for the intensity calibration of a broad spectrum, whereby many spectra are spliced together following repeated rotation of the grating. [63] The spectra of interest are produced by a white lamp and a fluorescence standard. Wavelength calibration using krypton was applied in advance. It was observed that the errors in the wavelength calibration follow a 'quadratic trend', although no further investigation is applied since this has little impact on the accuracy of the intensity calibration.

Martinsen et al. developed a protocol to calibrate a spectrometer with poor resolution [70] by using a filter to sequentially isolate single peaks in the wavelength reference, followed by calibration based on the recorded peaks. Of particular interest in the context of our work, is the use of a 'constrained cubic' polynomial for wavelength calibration, whereby the relationship between wavelength and pixel position is assumed to be predominantly linear with the residual term described by a weak third order polynomial. In Ref [42] a fifth order polynomial was used to relate wavelength and camera pixel for a neon–argon-lamp as part of a wavelength/wavenumber/intensity calibration routine for Raman spectra. To the best of our knowledge, this is the only instance of a polynomial order > 4 being used in a wavelength calibration routine.

Recently, there have been some efforts to improve the accuracy of wavelength calibration by first improving the quality of the reference spectrum in advance of calibration. [71, 72] This pre-processing includes denoising, stray-light removal [73], and deconvolution for the purpose of compensating for the spatial frequency response of the spectrometer, [74–76] as well as improved estimation of peak positions based on Voight or Lorenzian fitting. [43, 71, 77]

2.4.2 Wavelength calibration by modeling the physical system

Recently there has been interest in using a physical model of the optical path in the spectrometer for the purpose of wavelength calibration [78–84] and this is given special attention here since it is the core topic of Chapter 4. In the first such method [78] a wavelength calibration routine was developed based on modelling the optical system for the case of a Czerny Turner spectrograph using reflective concave mirrors. Similar to the method proposed in this paper, this method uses the diffraction equation to derive a relationship between the detector pixel position and the wavelength. A series expansion is applied to this equation, and only the first three terms are used. The resulting expression is a second order polynomial, the coefficients of which are defined in terms of the system parameters, including the grating angle, the deviation angle, the grating period, the focal length, the camera pixel size, and the tilt of camera. Assuming a known constant grating period, all of these parameters can be determined by a simple second order polynomial fit applied and examination of the resulting coefficients. Effectively, this second order polynomial fitting provides the basis for all future calibrations; using the parameters from the original second order fitting a single peak is sufficient to calibrate following thermal expansion, which affects the values of the focal length or deviation angle. Although this algorithm uses a model of the system, it is essentially a second order polynomial fitting method and is subject to the same errors that can result from polynomial fitting, in particular when the reference spectrum does not have lines that cover the full wavelength bandwidth of the spectrometer. It should be noted that the proposed algorithm also takes into account changes in the reference wavelength due to variation in the refractive index of the air taken from Ref. 83.

In Ref. 79 a wavelength calibration approach is proposed that also uses a physical model of the spectrometer, which replaces the need for polynomial-fitting described above with a brute-force search. The relationship between pixel position and wavelength is described as a function of the various system parameters including the grating angle. A brute-force search over these parameters is applied in order to find the best-fit for the recorded peaks from a reference neon lamp or similar. The key advantage of this

approach is the accuracy of the calibration outside the end-peaks in the reference since polynomial-fitting does will not extrapolate well in these bands, and also the ability of the method to be used with reference spectra containing only a small number of peaks. Liu et al. proposed the first instance of this approach in 2013, [79] and we provide a brief review of this work here, since it is most similar to the wavelength calibration algorithm proposed in this paper. A physical model of a Czerny-Turner spectrometer is used to derive the relationship between the three-dimensional coordinates of the camera port and the points at which the various wavelengths will come to focus. This physical model employs several system parameters relating to the four key elements in the spectrometer: (i) the angle of the collimating mirror, (ii) the angle of the grating, (iii) the angle and centre position of the imaging mirror, and (iv) angle and centre position of the detector. All of the aforementioned angles were taken to be one-dimensional, while the centre locations of the latter two elements were considered in two dimensions. Of these eight parameters, only four were included in the calibration algorithm as variables: the angle of the grating, and the angle and centre position of the detector. The remaining four parameters were assumed to be fixed and their values were measured. The calibration algorithm is based on a brute-force search in a predefined range over these four parameters, in order to find the set of parameters that provides the best fit for the recorded position-wavelength values.

Zhang et al. [80] were particularly interested in developing a model that could also account for a grating that was mounted on a sine-bar to achieve rotation. The model included several parameters relating to the mechanical function of the sine-bar. The authors identified six key parameters, which were functions of the sine-bar mechanical properties, as well as the grating period, the angle of deviation, and the centre of the detector. The wavelength calibration algorithm is based on solving a set of simultaneous equations that are derived from the physical model, in order to estimate the six key parameters. The authors state that this places a lower limit of five reference peaks for the algorithm to work. However, it should be noted that such an approach may be adversely affected by error in estimating a single peak position, whereas the iterative approach

described earlier [79] would likely be significantly more robust to errors in a single peak position value.

In Ref. 84 the authors use the diffraction equation to derive a single equation to relate pixel position and wavelength for a lens based reflection spectrometer. They identify three coefficients in this equation, each composed of two or more system parameters including focal length, deviation angle, grating angle, camera centre and pixel size. Solving for these three unknowns requires only three spectral lines from a mercury lamp. The authors report superior results compared with first, second, and third order polynomial fitting as well as two trigonometric methods. They report a 'standard error' (which is similar in definition to RMSE except that it includes the number of coefficients used in the model in its definition) of 0.05nm.

In Ref. 81, 82 a number of wavelength calibration algorithms are developed based on the physical model. This model accounts for all of the system parameters that are investigated by other researchers, [79, 80] but also accounts for tilt of the detector both horizontally and vertically, as well as accounting for displacement of the input irradiance vertically along the slit. As for other papers, the diffraction equation is used as the basis for deriving a model for the physical system. The algorithm fits up to nine system parameters to this equation although the method of fitting is not discussed. This is reduced to eight when the grating angle is known precisely using an optical encoder. With this encoder, the wavelength accuracy is reported to be 0.005nm and 0.025nm. It should be noted that the proposed algorithm also takes into account changes in the reference wavelength due to variation in the refractive index of the air according to Ref. 85.

We also note that modelling the physical system has also previously been considered for Echelle spectrometers. [86–88]

2.5 Wavenumber Calibration: Background

We note that the review that follows in this section (and the Appendix that immediately proceeds it) is included as part of a paper, *Liu, Dongyue, and Bryan M. Hennelly. "Wavenumber Calibration by Modeling the Raman Spectrometer," submitted to Journal of Raman Spectroscopy, Sept 2022.*

In this section, we provide a detailed review of direct wavenumber calibration in the literature, and compare with the alternative approach of wavelength calibration followed by wavenumber conversion. In this review we have compared the various contributions over several important features including: the reference materials used, the number of peaks in the reference spectrum, the resolution of the systems, methods for sub-pixel interpolation, and calibration accuracy. We believe this is the first such review of its kind in the literature to date. We note that this review is included as part of a paper, *Liu, Dongyue, and Bryan M. Hennelly. "Wavenumber Calibration by Modelling the Raman Spectrometer." submitted to Journal of Raman Spectroscopy Sept 2022*, which is currently under review.

There are two approaches to wavenumber calibration in Raman spectroscopy: (i) Initial wavelength calibration using an atomic emission spectrum such as neon (for a comprehensive review of wavelength calibration we refer the reader to the Ref. 89), followed by wavenumber conversion making use of the excitation laser wavelength in this calculation; (ii) Direct wavenumber calibration using a Raman reference standard such as acetaminophen or indene. Both approaches appear to be equally represented in the literature and both require the use of a reference standard with a spectrum containing several well defined narrow peaks, which are subject to polynomial fitting. In Table 2.1, a non-exhaustive literature review is provided of both approaches. This table compares the different reference standards, polynomial orders, sub-pixel interpolation methods, and accuracy of both approaches. To save space, the footnote of Table 2.1 are given in Appendix 2.6.1.

In the first publication to examine a wavenumber calibration protocol, [59] Hamaguchi et al. reported that wavelength calibration followed by wavenumber conversion

provided higher accuracy. However, our review of the literature reveals that both approaches provide similar levels of absolute error (accuracy) with a limit of approximately 0-0.1 cm^{-1} and a standard deviation (or precision) of 0-0.1 cm^{-1}. Direct wavenumber calibration is a single step approach with the advantage of not requiring *a priori* knowledge of the laser excitation wavelength, which can be difficult to record using the Raman system if the spectrograph grating is configured to avoid this wavelength, or if long-pass or notch filters remove it. Rotating the grating to record the laser also presents problems in that wavelength calibration must be repeatedly applied following rotation. In such cases, a separate spectrometer is sometimes used to record the laser wavelength. Alternatively, the laser wavelength can be estimated by measuring the Raman wavenumber shift of Raman reference standards such as cyclohexane. [68, 90] Error in the laser wavelength measurement is problematic; for example, in Ref. 91, the authors report an uncertainty of ±0.005 nm in the laser wavelength introduces an error of 0.18 cm^{-1} in the corresponding relative wavenumbers, at 532 nm excitation. Given the similar accuracy of the two approaches and the additional difficulties in the two step wavenumber conversion approach, it can be taken that direct wavenumber calibration is preferable.

Typically the wavenumber conversion approach uses neon as the reference material. However, the reference standards that have been investigated for direct wavenumber calibration are far more varied [101] as seen in Table 2.1. In Ref. 69 the authors define the characteristics of a suitable wavenumber reference standard. In summary, the spectrum must contain a large number of sharp peaks (of known wavenumber position) that extend over the full range of the recorded bandwidth; the accuracy of the calibration outside of the end peaks drops significantly. Since no single standard will provide this, it is common to see several reference materials combined in a calibration protocol. [39, 97] The sampling considerations of the standards is also important and should resemble those of the samples to avoid any change in the recording conditions. Gaseous samples can be troublesome as these require specialised containers, although these samples typically provide the sharpest peaks. Also important is the long term stability of the sample; photo-degradation of the sample or chemical change over time can result

Ref	Peak	Poly-nomial order	Reference material	Sub-pixel interpolation	Resolution (cm^{-1})	Calibration accuracy (cm^{-1})
Wavelength Calibration followed by Wavenumber Conversion						
59	21	3-5	neon + laser wavelength	-	-	SD:0.2-0.4 AE:0.2-0.5
68	39/16[†1]	5	argon, neon + laser wavelength[†1]	8×FFT	5-7	$\overline{\text{MAE}}$:0.85 $\overline{\text{SD}}$:0.257[†2]
92	5	1	neon + laser wavelength	Apodization + 4×FFT + Gaussian fit	10	MAE:1 $\overline{\text{SD}}$:0.2
93	50	1[†3]	thorium, neon + laser wavelength	Apodization + -×FFT + 4th order polynomial fit	4	$\overline{\text{SD}}$:0.04-0.06
90	59/21[†1]	1	neon + laser wavelength[†1]	Apodization + -×FFT + polynomial fit	-	RMSE:0.4 $\overline{\text{SD}}$:0.05
94	9	-[†4]	neon + laser wavelength	Gaussian fit	-	-
95	28	3	neon + laser wavelength[†5]	Gaussian fit	1.8	$\overline{\text{SD}}$:0.6
78	15	2	neon, mercury + laser wavelength	8×FFT + log-normal function fit	0.87-1.4	$\overline{\text{SD}}$:0.1
96	18/11[†6]	1	neon + laser wavelength[†6]	-×FFT + polynomial fit	-	SD:0.24/0.04[†6]
Direct Wavenumber Calibration						
59	17	-	indene	-	-	-
93	5	1	acetonitrile	8×FFT + 4th order polynomial fit	4	-
69	17	1-3	indene	8×FFT	-	MAE:(1:0.25, 2:0.2,3:0.16)[†7]
39	67	4	4-acetamidophenol, benzene, polystyrenel, acetonitrile	64×FFT + 2nd order polynomial fit	4	AE:0-0.1 SD:0.01-0.21[†8]
97	45	3	H_2, HD, D_2, O_2	Gaussian, Lorentzian, or Voigt fit	3.5	MAE:0.08-0.11[†9]
91	20	1	CO_2, CO, N_2O, H_2O, HC1	4th order polynomial fit	0.06	MAE:0.003-1 $\overline{\text{SD}}$:0.002
3	15	4	4-acetamidophenol	3rd order polynomial fit	2.5	SD: 0.0374-0.6542[†10]
98	8	2-4	indene	-	-	-
49	19	2-4	4-acetamidophenol	-	4.36-12.03	RMSE: 0.52-3.56[†11]
99	13	3-5	4-acetamidophenol	Gaussian or Lorentz fit	-	-
45	3	-	polystyrene, benzonitrile, cyclohexane[†12]	-	-	nIQR:0.74[†13]
44	10/5/5	3	4-acetamidophenol, cycohexane, polystyrene	Gaussian fit	1-20	MAE:0.45-1.93 SD:0.24-2.16[†14]
100	-	3	acetone, sodium perchlorate, potassium perchlorate, acetonitrile, military C-4.	Gaussian fit	2-5	AE:0-30

Table 2.1: Non-exhaustive literature review of wavenumber calibration for Raman spectrometers. The two approaches of wavelength calibration followed by wavenumber conversion, and direct wavenumber calibration are compared in terms of: reference materials used, number of peaks, polynomial order, method for sub-pixel interpolation to identify the peak positions on the detector with high accuracy, resolution of the systems, and the reported accuracy/precision. For the latter, the various metrics given here as abbreviations are defined later in Section 5.5.2. To save space, footnotes are provided in Appendix 2.6.1.

in errors. One final note on the selection of wavenumber reference standards is on the use of a single peak standard such as Silicon. Such a standard is sometimes used to account for a constant offset in the wavenumber axis due to small daily changes in the system/environment. It has recently been shown that this approach results in significant error; [45] it is notable that the European Pharmacopoeia describes the requirement of a minimum of three wavenumber shifts covering the working range of the instrument. [102]

A common feature in wavenumber calibration using either of the approaches is the identification of peak position with sub-pixel accuracy, as highlighted in the 'sub-pixel interpolation' column in Table 2.1. Any error in identifying these peak positions will consequently result in a calibration error. A desirable characteristic of the wavelength/wavenumber standards in the peaks are narrow; however, this presents a problem for accurately identifying the position of the centre of peak if the sampling interval of the detector is of similar width or larger than the peak width. An early approach involved the recording of many spectra with very small shifting of the spectrum between captures. [103] To overcome the problem for a single spectrum, many authors use a first step of apodization, [90, 92, 93] whereby the raw spectrum is subject to a discrete Fourier transform (DFT), followed by multiplication with a Gaussian function or similar and an inverse DFT. The result is a convolution of the raw spectrum with a symmetrical blurring function, which broadens the shape of the peaks such that the centre can be more easily identified. The DFT is implemented with computational-efficiency using the fast Fourier transform (FFT) algorithm. An additional step of interpolation is commonly applied, whereby the DFT of the spectrum is zeropadded to increase the number of samples n-fold, which provides n times interpolation. In the table this is denoted as $n\times$FFT. Finally, in order to identify the position of the peak with a resolution smaller than the sampling interval, the peaks are fitted using a Gaussian, Lorentzian, or Voigt function; [44, 92, 94, 95, 97, 99, 100] and the centre of the fitted function is taken to be the peak position. Alternatively, polynomial fitting [3, 39, 90, 91, 93, 96] can be applied to the few samples around the maximum peak samples; the zero valued derivative of

the polynomial in this region reveals the local maximum, which is taken as the peak position.

Once the position on the detector of each peak in the reference spectrum have been identified with sub-pixel accuracy, these values are then subjected to polynomial fitting with respect to their known wavelength or wavenumber values. In this way the wavelength or wavenumber axis for the spectrum is determined. For the former, the wavelength axis is then converted to wavenumber values making use of the laser wavelength. Although the first paper to investigate wavenumber calibration employed a polynomial order of 5 (necessitated by the presence of high levels of distortion in the imaging system), a low order polynomial is generally preferred in the literature for both the wavenumber conversion and direct wavenumber calibration approaches, as seen in Table 2.1. In Ref. 98, the authors compared the use of 2^{nd}, 3^{rd}, and 4^{th} order polynomial fits. They demonstrate that 4^{th} order fits give significantly worse results in comparison to 2^{nd} and 3^{rd} order fits. Their data also indicates that better fits were obtained with a quadratic equation for three (413.1, 487.9, and 514.5 nm) of the five excitation wavelengths examined. Interestingly, it has been shown that a quadratic expression is sufficient to model the relationship between wavelength and a CCD detector based on the physics of dispersion from a diffraction grating and imaging in the presence of low distortion; the coefficients of the second order polynomial can be defined in terms of the system parameters. [78] Recently, we have shown that the non-linearity of this relationship varies considerably as a function of grating period and focal length. [89] It can be expected that the non-linearity of the wavenumber axis will be more pronounced given the non-linear relationship between wavelength and wavenumber, and therefore, a polynomial order > 2 may be mentioned. We examine this point in more detail in Section 5.2. It is notable that recent large scale cross-instrument investigations have preferred 3^{rd} order polynomials for direct wavenumber calibration. [44, 49, 100]. Some authors have employed linear interpolation between adjacent reference peaks as an alternative to polynomial fitting with high levels of accuracy. [93, 96] A key point that is often overlooked is the accuracy of the points used in the polynomial fitting; while

great attention is given to finding the position of the peaks on the detector with sub-pixel accuracy, the accuracy of the true wavelength/wavenumber values associated with these peaks is rarely discussed. For the case of wavelength calibration, the wavelength of the peaks in atomic emission spectra is subject to change depending on environmental conditions; spectral line positions for the reference lamp should be corrected to account for the refractive index of air. [83, 85] Rarer still is any consideration of the shift in wavenumber standards due to variation in temperature. It is known, for example, that the peaks in the cyclohexane spectrum, a commonly employed wavenumber reference standard, will change both in terms of area and offset. [104] Furthermore, the change in each peak is different and depends on the nature of the vibrational/rotational mode and also on the level of depolarisation associated with that peak; some peaks were found to vary by 25 times more than others. [104] It is difficult to account for such changes in a wavenumber reference spectrum; however, as noted in Ref. 104, for the case of cyclohexane, the temperature dependence of the Raman band position is small enough to be ignored for the calibration requirements of most applications (for example, the 802 cm^{-1} band shifts 0.12 cm^{-1} for a 10 degree temperature change). Given the consistency of measurement in the literature to date, it can be expected that other wavenumber standards will have similar behavior. One final point of note on polynomial fitting, is the location of the outermost peaks in the reference spectrum. It is very important that peaks exists close the ends of the recorded band; this is necessary to avoid high levels of error with polynomial fitting in bands outside of the range of the reference peaks. This has necessitated the combined use of several wavenumber reference standards in order to fully cover the bandwidth of the spectrometer. In this chapter, we develop a calibration method that is significantly more robust to this problem than polynomial fitting.

Any discussion of calibration accuracy of Raman instruments should begin with mention of the European Pharmacopoeia (Ph. Eur.), [102] which defines the tolerance for the wavenumber shift in the recording of several peaks of polystyrene, paracetamol, and cyclohexane to be ±1.5 and ±2.0 for bench-top and handheld instruments respec-

tively. The accuracy of the various calibration protocols that have been reported in the literature to date vary considerably, as seen in Table 2.1. Given the different metrics for 'accuracy' and 'precision' that have been applied to date, it can be difficult to directly compare many of these references. We refer the reader to Section 5.5.2 for explicit definition of these various metrics. It is notable for many papers, only the standard deviation is reported as a measure of precision while the accuracy of the calibration with respect to the known reference peaks is not reported. The reason for this is that the wavenumber shifts of these standards were not yet defined with an agreed upon accuracy. Assuming an accurate sub-pixel interpolation method is used, three factors appear to determine the calibration accuracy: (i) the resolution of the system as evidenced by the high accuracy for FT-Raman calibration; [91, 93] (ii) the availability of a high number of known reference peaks over the band of interest; (iii) the sharpness of the peaks; and (iv) the accuracy of their 'known' wavenumber shifts. These points are emphasised by the most accurate calibration protocols reported in the literature to date in Ref. 39 and Ref. 97. In Ref. 39, the authors use 67 peaks from a combined reference standard and evaluate accuracy over several months. The standard deviation of the wavenumber shifts is reported to be in the range 0.01-0.21 cm^{-1}, which is the lowest reported to date for the specific reference standards used in the study including the values reported by ASTM. [105] The absolute error with respect to the ASTM wavenumber shifts was zero for almost all of the peaks reported; however, this should take into account the limted accuracy for these standards reported by ASTM of 0.1 cm^{-1}. This study highlights the importance of using a large number of peaks for calibration, but also the limitation in terms of accuracy afforded by most of the calibration standards published by ASTM. In Ref. 97, gaseous reference standards are combined, which have extremely sharp peaks and for which the wavenumber shift are known down to an accuracy in the order of 0.001cm^{-1}. [106–108] In total, 45 peaks are used and the mean absolute error is reported as 0.08-0.11 cm^{-1}. However, the authors have estimated error by taking the average of the residual error in wavenumber of the reference peaks and the residual pixel error subsequently converted into wavenumber error. If the authors had used only residual

error in the wavenumber shift of the reference peaks to define absolute error, which is the typical approach in the literature, their reported mean absolute error would be 0.001652-0.000818 cm^{-1}, which would represent the most accurate calibration reported thus far in the literature, albeit for a single spectrum.

In this chapter, we propose a new direct wavenumber calibration protocol that advances on the background material reviewed in this section. More specifically, we propose an alternative to the third-order polynomial fitting step used in existing direct wavenumber calibration protocols; although the accuracy and precision that have been reported by the state-of-the-art [39, 97] is very high, we propose to augment these protocols such that they are more accurate in bands outside of the outermost peaks; we note that all of the literature to date does not report the accuracy in these bands. The method proposed here may also enable few peaks to be used and therefore reduce the number of materials used in the composite reference standard. We will demonstrate that our method has slightly higher accuracy than second and third-order fitting inside the band of the reference peaks, and significantly higher accuracy outside of this band. The proposed method is based on using a physical model of the Raman spectrometer to derive a relationship between wavenumber and detector pixel. The use of a physical model of the spectrometer has previously been used in Raman wavenumber calibration; [78] however, in that case the author relates wavelength to pixel and the method can therefore not be classified as a direct wavenumber protocol. Furthermore, the author defines the wavelength/pixel relationship to be governed by a second order polynomial, the coefficients of which are defined by the parameters of the optical system. This is done only once and these parameters are fixed for future re-calibration. In this chapter, we do not impose such a limitation and we assume a general rotating grating, which can be transmitting or reflecting and which requires frequent calibration.

2.6 Appendix

2.6.1 More details on Table 2.1

Footnotes for Table 2.1 are provided here: [†1]: The laser wavelength is estimated using the wavenumber shift of the cyclohexane and toluene peaks. The first number in column 2 relates to peaks from the atomic emission lamp, the second number relates to combined cyclohexane and toluene peaks used to estimate laser wavelength; [†2]: The values are calculated from Ref68 Table vii; [†3] Different first order polynomials were used to connect each set of adjacent reference peaks; [†4]: Polynomial fitting is not used; A physical model of the system is developed relating wavelength to pixel position, which is used to determine the wavelength axis; [†5]: The laser wavelength is estimated by measuring the wavenumber shift associated with the rotational lines of hydrogen; [†6] 18 neon peaks were used for wavelength calibration, and the laser wavelength is estimated using the wavenumber shift of 11 cyclohexane peaks. The SD values relate to a single peak of cycolhexane. The smaller value is the result of several hundred separate calibrations averaged together where the grating is slightly moved between calibrations. [†7]: SD values are provided as follows: indene fingerprint region SD:0.4-1 and CH-band SD:0.5-2. The MAE results given in the table relate to a simulation of calibration taken from Table I and Table II in Ref 69 using polynomials of orders 1, 2, and 3. [†8]: Results for Bis(MSB) and Naphthalene are compared with ASTM E1840-96 [2]. Mean wavenumber positions over 60 measurements are identical for all peaks except four, each of which had a difference of 0.1 cm^{-1}; SD for Bis peaks are 0.02-0.21 and for Naphthalene are 0.01-0.13. [†9]: The maximal error of residuals in a single (combined) spectrum, 3-sigma, is actually reported as 0.24, and 0.34 for parallel and perpendicular polarisation, respectively. The authors have estimated error by taking the average of the residual error in wavenumber of the reference peaks after third order polynomial fitting (this is the standard approach by other authors) and the residual pixel error, which is subsequently converted into wavenumber error. The latter is much larger than the former; if the authors had used the typical approach their reported errors would be MAE:0.001652 and 0.000818, which

would represent the most accurate calibration reported thus far in the literature. [†10]: SD of 15 peaks from 4-acetamedophenol over 100 spectra ranged from 0.0374-0.5333; SD of 15 polymer peaks over 100 spectra ranged from 0.0769-0.6542. [†11]: Three devices, and three polynomial orders (2,3,4) were evaluated: Device 1:(3.56,3.49,3.47); device 2(0.98,0.92,0.91); device 3(0.76,0.60,0.52). [†12]: For three chemicals only a single peak was analysed. These are 1001.3cm^{-1} for polystyrene, 1001.1cm^{-1} for benzonitrile, 802.0cm^{-1} for cyclohexane; [†13]: The normalised interquartile range (nIQR) is given as an estimate of the standard deviation; [†14]: Results are reported separately for a single spectrum of the three standards: paracetmol(MAE:0.45,SD:0.24); cyclohexane(MAE:1.93,sd2.16); polystyrene(MAE:1.31,sd 1.52).

Chapter 3

Wavenumber calibration using a Polymer Reference standard

The work in this chapter has been published in the following reference: *Liu, Dongyue, Hugh J. Byrne, Luke O'Neill, and Bryan Hennelly. "Investigation of wavenumber calibration for Raman spectroscopy using a polymer reference." In Optical Sensing and Detection V, vol. 10680, pp. 486-497. SPIE, 2018.* with the following abstract:

"Raman spectroscopy is an optical technique that can be used to evaluate the biomolecular composition of tissue and cell samples in a real-time and non-invasive manner. Subtle differences between datasets of spectra obtained from related cell groups can be identified using multivariate statistical algorithms. Such techniques are highly sensitive to small errors, however, and, therefore, the classification sensitivity of Raman spectroscopy can be significantly impacted by miscalibration of the optical system due to small misalignments of the optical elements and/or variation in ambient temperature. Wavenumber calibration is often achieved by recording the spectrum from a wavenumber reference standard, such as 4-acetamidophenol or benzene, which contains numerous sharp peaks in the fingerprint region. Here, we investigate a commercial polymer slide as a wavenumber reference standard for the calibration of Raman spectra. The Raman spectrum of this slide contains numerous sharp peaks in the fingerprint region. Unlike many other reference standards, the polymer slide is non-hazardous, has an indefinite

lifetime, and is designed in the shape of a glass slide used for microscopy. We evaluate this reference in terms of accuracy and repeatability, and we compare with the established 4-Acetamidophenol wavenumber reference."

3.1 Introduction

A number of solutions that perform wavenumber calibration already exist. [109] One solution uses two steps: (i) a known spectrum from a Neon lamp (or similar) allows wavelength calibration of the spectrometer to be performed, [39] followed by (ii) subsequent wavenumber calibration, which is performed using a known standard such as silicon, which produces a well-known sharp peak at 520 cm^{-1}. A second solution, which uses a single step, is to employ a wavenumber reference standard such as indene, [98, 110] cyclohexane, [62, 111, 112] benzene, [113] and benzonitrile. [114] These standards contain numerous sharp peaks in their Raman spectra at well-known wavenumber locations. A polynomial function can be fitted to these peak positions in order to calibrate the entire wavenumber axis. Often, many wavenumber standards are used at the same time in order to increase the range of, and improve the accuracy of, the calibration.

All of the chemical standards that we reviewed in the literature are hazardous to human health and must be handled in a controlled manner. In this chapter, we propose a new wavenumber reference in the form of a commercial polymer slide that is designed for life science applications, which is inexpensive, safe to handle, and chemically stable over time. In Section 2.3, the various sources of error in a Raman spectrometer are discussed and, in Section 3.2, the traditional approach of wavenumber calibration using a wavenumber reference is reviewed, specifically for the case of 4-Acetamidophenol, a commonly used wavenumber reference. In Section 3.3, the peak positions of the polymer are measured and the stability of these peak positions is compared with that of 4-Acetamidophenol. In Section 3.4, a brief conclusion is offered.

3.2 Wavenumber calibration using a wavenumber reference

The purpose of this section is two-fold. Firstly, the protocol used for wavenumber calibration using a wavenumber reference is described in detail and experimental results are shown describing each step using a sample of 4-Acetamidophenol, a commonly used wavenumber reference standard. Secondly, the newly calibrated instrument is used to record the (calibrated) spectrum from a commercial polymer slide that is commonly used for life science purposes (μ-Slide I Luer, Ibidi GmbH, Munich, Germany). [115] This slide is designed with a flow channel for imaging adherent cells under flow conditions as well as 3D cell culture. The base of the slide is made of a transparent polymer with coverslip thickness to facilitate imaging with an inverted microscope. The properties of autofluorescence, birefringence, and the refractive index of the Ibidi polymer coverslip are similar to those of glass, allowing for the use of all kinds of objective lenses including oil immersion; the specific chemical structure of the polymer material is proprietary, and could not be ascertained. In Section 3.3, we explore the potential of this polymer slide to be used as a wavenumber reference by analysing the stability of the peaks in the polymers Raman spectrum and comparing these results to those obtained from 4-Acetamidophenol.

The protocol for using a wavenumber reference to perform wavenumber calibration has been developed in many other references. [39–42, 116] The steps involve:

(i) Recording the spectrum from a sample with a pure chemical that has a known Raman spectrum containing a number of sharp peaks at well-defined wavenumber positions.

(ii) Obtaining the precise wavenumber positions from a reliable source such as the American Society for Testing Materials (ASTM) or the National Institute of Standards and Technology.

(iii) Recording the sample (pixel) position of each of these peaks in the spectrum, and pairing these with their respective wavenumber positions to provide a set of

two-dimensional coordinates in the form (pixel, wavenumber).

(iv) Applying a polynomial fit to these coordinates using the method of classical least squares. The resulting polynomial provides the relationship between every pixel and the corresponding wavenumber.

Further detail on these steps can be found in Ref. 39. The number of peaks that are needed to perform accurate calibration has been a subject of interest in the literature, as has the order of the polynomial that should be used in the Step (iv). [39–42, 116] The accuracy can be shown to be dependent on the number and position. The peaks and a large number of peaks covering as wide an area of the spectrum should be used. In order to increase the number and range of peaks used in the calibration, some researchers have used multiple chemical references in a single calibration. [39] The order of the polynomial should not be so large as to result in overfitting but not so small to result in under fitting; a polynomial order of four has been shown to perform well. [39] The pixel positions that are recorded in Step (iii) can be obtained with sub-pixel accuracy using a process of interpolation as proposed in Ref. 39. In the experiment that follows, we perform cubic-spline interpolation in the area of each peak in order to accurately identify the position of each peak.

Spectra were recorded using a custom built confocal Raman microspectrometer operating with a 532nm laser (150mW, Torus; Laser Quantum, Cheshire, UK), 10x microscope objective (10 / 0.3 Olympus MPlanFl; Olympus Corporation, Japan), and 100 μm confocal aperture. Raman scattered photons are collected with a spectrograph (Shamrock 500; Andor Technology, UK) operating with a 1000 lines/mm grating (spectral resolution of 2.5 cm^{-1} at the centre), and a cooled CCD camera (Newton 920; Andor Technology, UK) operating at −80°C. More details on this optical system are available in Ref. [23]. The chemical used to calibrate our instrument is 4-Acetamidophenol, a commonly used wavenumber reference standard. A pure sample of this chemical was obtained from a commercial source (Sigma Aldrich, Wicklow, Ireland) and the wavenumber positions of the peaks were obtained from an international standards organisation (ASTM International, Pennsylvania, US), which are shown below in Table 3.1.

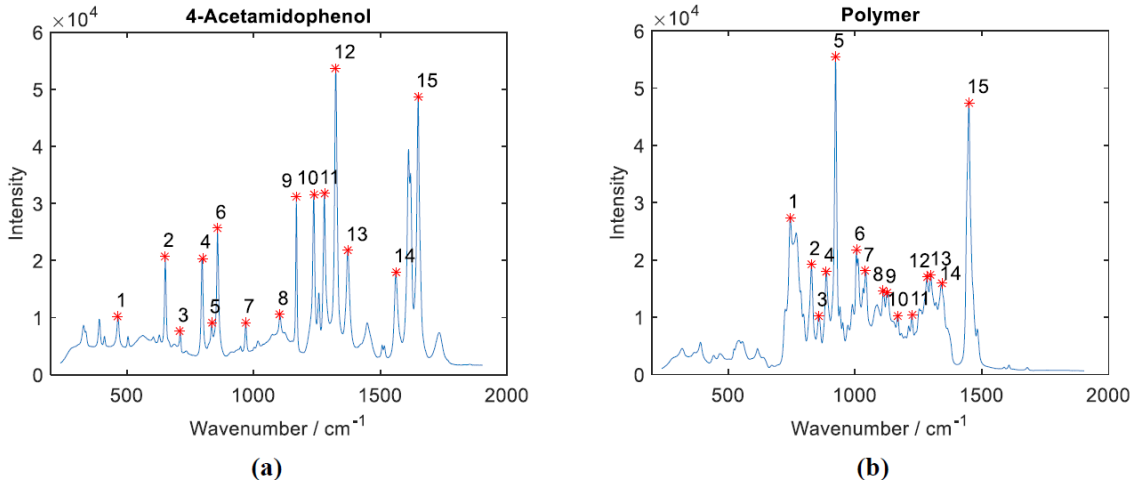


Figure 3.1: (a) Raman spectrum of the 4-Acetamidophenol spectrum. (b) Raman spectrum of the polymer sample. For both cases, the wavenumber axis has been calibrated using the calibration protocol described in this section, applied to the peaks positions of the 4-Acetamidophenol spectrum

Five spectra, each with an acquisition time of 16s, were recorded from the sample (acetaminophen), contained in a vial with a base made from a Raman grade Calcium Fluoride (Crystran, UK) coverslip with a thickness of 200μm. This acquisition time was just less than that which would cause saturation of the camera. These five spectra were averaged to produce a low noise Raman spectrum with a total acquisition time of 80s, which completes Step (i). This average spectrum is shown in Fig. 3.1 (a). The sharpest 15 peaks were selected for use in the calibration process and these are numbered from 1-15 in the figure. The wavenumber position of each of these peaks was obtained from Table 3.1, completing Step (ii). In Fig. 3.2 (a), the fifteen coordinates that result from Step (iii) as well as the fourth order polynomial that was fitted to these coordinates in Step (iv) are both shown. The positions of the peaks is determined with sub-pixel accuracy using cubic spline interpolation. [115] The third order polynomial that is associated with each of the fifteen peaks is isolated and the derivative of this function provides a solution for an accurate (sub-pixel) position of each peak. This process is illustrated in Fig. 3.2 (b). These values of the pixel position of the peaks are used to define the coordinates used for polynomial fitting, shown in Fig. 3.2 (a). The calibrated wavenumber axis, which is obtained by relating each pixel position to its corresponding wavenumber value using this polynomial function, is shown in Fig. 3.1 (a).

Peak Number	4-Acetamidophenol Reference (cm^{-1})	Polymer Spectrum (cm^{-1})
1	465.1	743.5267
2	651.6	827.9567
3	710.8	857.9544
4	797.2	886.563
5	834.5	923.0671
6	857.9	1005.7366
7	968.7	1041.7295
8	1105.5	1111.6935
9	1168.5	1127.9242
10	1236.8	1168.3634
11	1278.5	1224.6504
12	1323.9	1283.9253
13	1371.5	1297.9906
14	1561.6	1340.9552
15	1648.4	1449.3456

Table 3.1: The reference table of spectral peak positions for a sample of 4-Acetamidophenol (ASTM E1840-96) and the (calibrated) spectral peak positions of the polymer sample.

Immediately following calibration of the system, as described above, five spectra, each with an acquisition time of 16s, were recorded from the polymer slide. In this case no sample container was necessary; the slide has the same dimensions as a traditional glass slide used in microscopy (7cm × 2.5cm × 1mm), and can be easily placed on a microscope translation stage. Once again, these five spectra were averaged to produce a low noise Raman spectrum with a total acquisition time of 80s.

The fifteen sharpest peaks were selected for further inspection, and once again a process of cubic spline interpolation was used to identify the position of the peaks with sub pixel accuracy. The positions were then related to their corresponding wavenumbers using the polynomial function returned from the calibration process already described. The wavenumber position of each of the fifteen peaks is also shown in Table 3.1 . Interestingly, the spectral intensity of the polymer is approximately the same as that from the 4-Acetamidophenol for the same acquisition time, indicating that, if this sample were to be used as a reference, similar acquisition times could be used. In terms of its applicability as a wavenumber reference, it can be seen that the polymer spectrum contains a large number of sharp peaks, albeit over a smaller range of the wavenumber

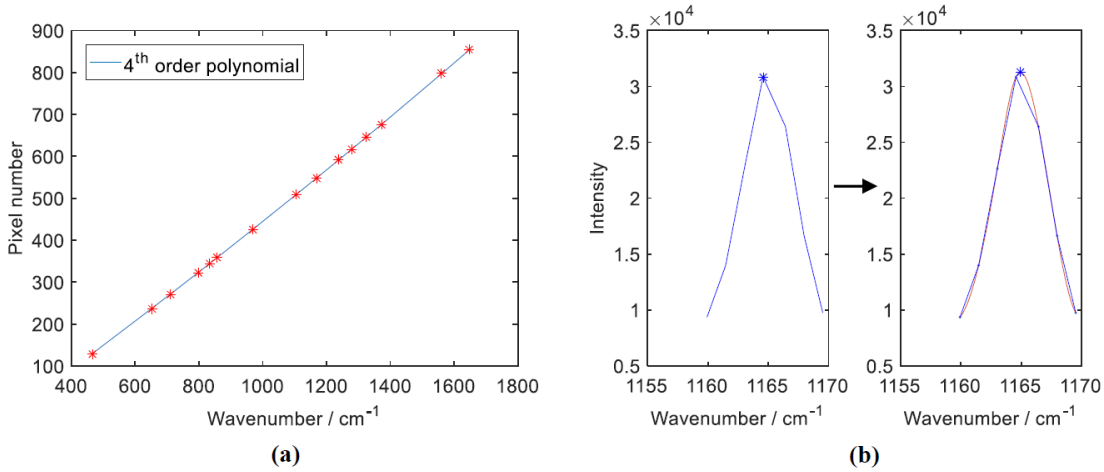


Figure 3.2: (a) Plot of the fifteen coordinates obtained by Step (iii) of the calibration protocol using 4-Acetamidophenol, as well as the polynomial fit from Step (iv); (b) Illustration of the interpolation process used to obtain sub-pixel accuracy for one peak position. The blue asterisk indicates the peak position with and without interpolation.

axis. In Section 3.3, these peaks are investigated further; specifically, we investigate the stability of the peaks positions across different sets of measurements.

3.3 Evaluation of polymer wavenumber reference

In order to evaluate the potential for the polymer slide to be used as a wavenumber reference standard, it is necessary to investigate the stability of the various peaks in terms of wavenumber position. With any peak measurement in Raman spectroscopy, there can be expected small deviation in the position of that peak across a set of measurements. The reason for the instability in peak position is due to the presence of noise, which can affect attempts to accurately measure a peak position. The noise in a Raman spectrum is primarily comprised of shot noise, dark current, read noise; the latter two noise sources originate from the camera but are insignificant compared to the shot noise generated by the types of samples used as wavenumber references. Shot noise is a time-dependent noise contribution that originates from the signal itself. Shot noise [117] is the name given to inconsistent levels of irradiance that are incident on a pixel over a given time, t. Irradiance per pixel, i, is typically modeled as a Poisson distribution. The signal to noise ratio (SNR) is given by $it/(it)^{0.5}$, which is the signal intensity divided by the standard deviation of the shot noise. It is clear that the SNR increases as a function of the square root of the acquisition time. It is, therefore, important to use a sufficiently long exposure time when recording a reference spectrum. The sharpness of the peak is also a consideration, since an estimation of peak position using the interpolation approach outlined in the previous section will make use of neighboring samples and will, therefore, include their noise contributions in the estimation.

An experiment was conducted to measure the wavenumber stability of the fifteen 4-Acetamidophenol peaks listed in Table 3.1 as well as for the fifteen polymer peaks. The position of each peak was measured across 100 consecutive recordings, for acquisition times of 1s, 2s, and 4s. In order to study the effect of reducing the SNR by successively halving the acquisition time. The position of each peak across the sequence of recordings is shown in Fig. 3.3 (a) for 4-Acetamidophenol and in Fig. 3.3 (b) for the polymer slide.

A comparison of the peaks that appear in the spectrum of 4-Acetamidophenol, shown in Fig. 3.1 (a), with the variation in the corresponding peak positions shown in

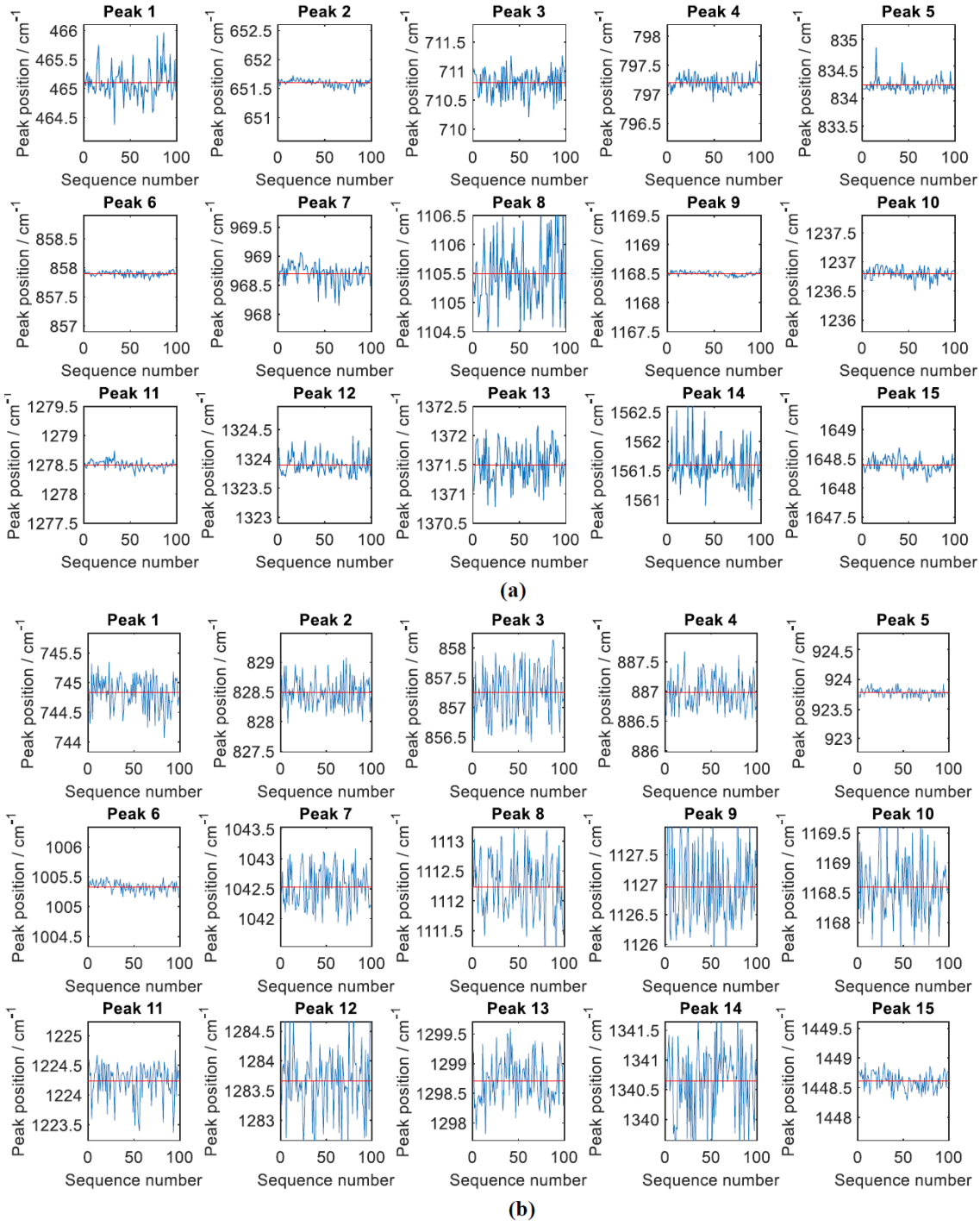


Figure 3.3: (a) The peak position of the fifteen peaks of 4-Acetamidophenol that are listed in Table 3.1 for a sequence of 100 spectra with 4s acquisition time. The red line indicates the mean position of the peak. In all cases the vertical axis has a range of $2cm^{-1}$; (b) The peak position of the fifteen peaks of the polymer slide that are listed in Table 3.1 , also for a sequence of 100 spectra with 4s acquisition time. The red line indicates the mean position of the peak. In all cases the vertical axis has a range of $2cm^{-1}$.

Peak number	Standard deviation of peak positions for 4-Acetamidophenol			Standard deviation of peak positions for polymer slide		
	1 second	2 seconds	4 seconds	1 second	2 seconds	4 seconds
1	0.636	0.5704	0.2653	0.5992	0.3988	0.283
2	0.3745	0.0718	0.0562	0.5073	0.3221	0.2523
3	0.422	0.3405	0.2207	0.988	0.6128	0.4502
4	0.398	0.1755	0.1203	0.519	0.3971	0.2721
5	0.5891	0.3302	0.1653	0.1481	0.11	0.0769
6	0.2525	0.0731	0.0447	0.2056	0.117	0.0845
7	0.4598	0.3417	0.1719	0.6429	0.5208	0.342
8	0.91	0.7974	0.5333	1.0037	0.6943	0.5167
9	0.2918	0.0507	0.0374	1.2029	0.7591	0.6542
10	0.2961	0.1478	0.0983	1.0578	0.6561	0.5158
11	0.3817	0.108	0.0694	0.5709	0.4253	0.3002
12	0.4064	0.2349	0.1665	0.8866	0.6989	0.5604
13	0.6736	0.403	0.3238	0.7638	0.6362	0.3891
14	0.6036	0.4034	0.3475	0.9761	0.8536	0.6106
15	0.3693	0.1708	0.1154	0.2701	0.1751	0.1502
Mean	0.4710	0.2813	0.1824	0.6895	0.4918	0.3639

Table 3.2: It shows the standard deviation of the wavenumber position of each peak for both samples for three different acquisition times. The values for the '4 sec' columns are standard deviation of the functions shown in Fig. 3.3.

Fig. 3.3 (a) indicates that the height of the peak should not be taken as a measure of expected peak stability. Peaks 2, 6, and 9 are the three most stable peaks, and none of these are the highest peaks. It appears from a qualitative inspection that the narrower peaks are the most stable. The same wavenumber range (2 cm^{-1}) is used for all of the figures that appear in Fig. 3.3 (a) and it is clear there is significant variability in peak stability across the fifteen peaks. The equivalent results are shown for the polymer slide in Fig. 3.3 (a). In this case peaks 5, 6, and 15 are the three most stable. In general, however, it is clear that the variability of the 4-Acetamidophenol peak positions is approximately half that of the polymer peak positions. The standard deviation of each peak position shown in Fig. 3.3, is listed in Table 3.2 , as well as the standard deviations for the case of a 1s and 2s acquisition time.

To provide a visual comparison, the standard deviation of each peak position is represented in graphical format in Fig. 3.4. Fig.3.4 (a) shows the standard deviation of all fifteen peaks in the 4-Acetamidophenol spectrum across each dataset of 100 spectra.

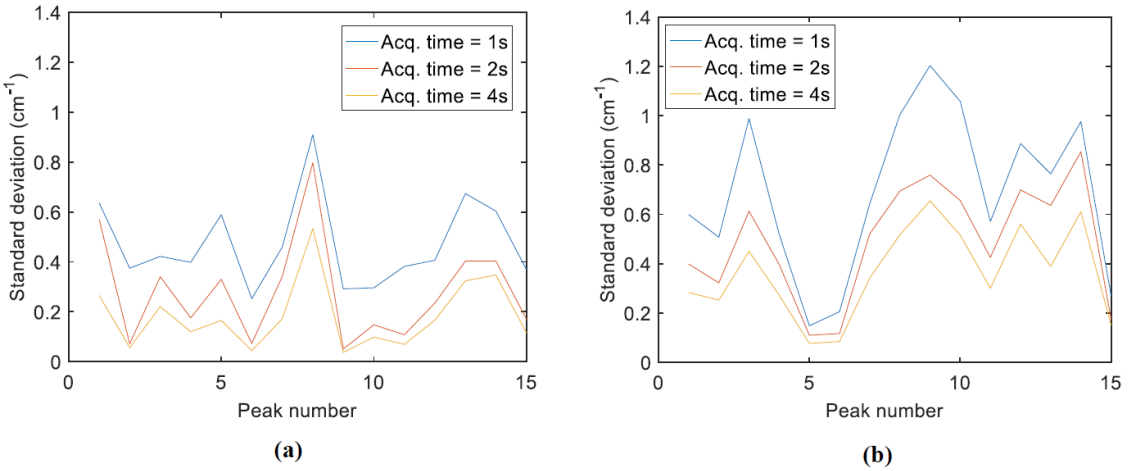


Figure 3.4: The standard deviation in wavenumber position for each of the fifteen peaks listed in Table 3.1 for three different acquisition times for (a) 4-Acetamidophenol and (b) the polymer slide.

The three datasets for the 1s, 2s, and 4s acquisition times are shown in different colours in the figure. The range of values is (0.25 cm^{-1} – 0.91 cm^{-1}), (0.05 cm^{-1} – 0.8 cm^{-1}), and (0.04 cm^{-1} – 0.53 cm^{-1}) for the 1s, 2s, and 4s datasets, respectively. As the acquisition time increases, the SNR increases with a square root relationship to time, and it can, therefore, be expected that there will be a related improvement in peak stability; the mean standard deviation of the peak position for the 4s case is 0.1824 cm^{-1}, which is 0.64 times the mean standard deviation for the 2s case, and 0.38 times the standard deviation for the 1s case. A similar trend can be seen for the polymer slide; Fig.3.4 (b) shows the standard deviation of all fifteen peaks in the polymer spectrum across each dataset of 100 spectra. The range of values is (0.21 cm^{-1} – 1.20 cm^{-1}), (0.11 cm^{-1} – 0.85 cm^{-1}), and (0.08 cm^{-1} – 0.61 cm^{-1}) for the 1s, 2s, and 4s datasets, respectively, which are similar to the values for 4-Acetamidophenol. A reduction in standard deviation is again observed for increased acquisition time; the mean standard deviation of peak position for the 4s case is 0.36 cm^{-1}, which is 0.73 times the mean standard deviation for the 2s case, and 0.52 times the standard deviation for the 1s case. The mean values of the standard deviation of the peaks in the polymer spectrum are approximately double the corresponding values for 4-Acetamidophenol for the 2s and 4s case.

3.4 Discussion

In this chapter, we propose a novel wavenumber reference for the calibration of Raman spectra. This material has a number of advantages over existing wavenumber reference materials. These slides are commercially available for life science applications and are manufactured to the same specification as a common glass slide used in microscopy. As such, the slides are ideal for placement on a microscope translation stage. Traditional reference materials are usually associated with health hazards and must be handled with care, and housed in sealed containers that use a window made of glass or a crystal that produces little Raman scattering, such as Quartz or Calcium Fluoride. The polymer reference material proposed here can be used without any of these considerations. The slide is inexpensive (<€10) and is chemically stable over time, unlike chemicals such as 4-Acetamidophenol, which will inevitably oxidise over an extended duration. The polymer slide appears to be robust to focused laser light and no melting was observed during our experiments using 150mW of a 532nm laser focused using a 10x/0.3 magnification objective. Melting was observed, however, using a 100x/0.9 magnification but stopped with a 50% reduction in laser power. It should be noted that a melting polymer material can irreparably damage the surface of a microscope objective, and care should be taken in this regard.

In Section 2.3 the various sources of error associated with miscalibration of a Raman spectrum were examined using a series of simulations. Specifically, we investigated an error in grating angle, lateral and rotational camera displacement, and laser wavelength instability. The impact of these errors on the wavenumber axis of a Raman spectrum was simulated using the diffraction grating equation and a simulation of the optical system within a Czerny-Turner spectrograph. It was clear that even small errors can lead to errors in the wavenumber axis of up to 100 cm^{-1}. Analysis of a Raman peak that was analysed over the course of 1000 recordings over a time period of approximately 3 hours, revealed a movement of the peak by almost 2 cm^{-1} which could result from a small change in ambient room temperature. It can be concluded that temperature control must be applied and frequent wavenumber calibration must be performed throughout

daily experiments.

In Section 3.2 the protocol for calibration of a Raman spectrum using a wavenumber reference was discussed in detail and the position of the fifteen most prominent peaks in the polymer spectrum was found following wavenumber calibration with 4-Acetamidophenol, a commonly used reference material, the Raman spectrum of which contains numerous sharp peaks throughout the fingerprint region. This calibration protocol includes a step to identify the pixel position of the given peak with sub-pixel accuracy using spline interpolation. The intensity of the polymer spectrum was found to be approximately the same as the intensity of the 4-Acetamidophenol spectrum, which indicates that similar acquisition times could be used for the polymer material if it is used as reference. The results in Section 3.3 indicate that on average the stability of the peak positions in the polymer spectrum is approximately half that of the peaks in the 4-Acetamidophenol spectrum; for a 4s acquisition the mean standard deviation of peak position is 0.18 cm^{-1} for 4-Acetamidophenol and 0.36 cm^{-1} for the polymer slide. Based on trends of peak stability for different signal to noise ratios, it can be concluded that these values will reduce significantly for a longer acquisition time, for example 20s. It was not possible to conduct the stability experiment in Section 3.3 for longer acquisition times because of temperature variability, which has been explained in Section 2.3.

The accuracy of the wavenumber calibration protocol that is used in Section 3.2 could be questioned due to the limited wavenumber resolution of the peak positions in the 4-Acetamidophenol spectrum that were provided by ASTM. These peak positions are accurate to only 0.1 cm^{-1} and it can be expected that a more accurate calibration of the polymer spectrum could be obtained given a more accurate set of peak positions for the reference material that is used to calibrate the polymer spectrum. Furthermore, the spectrograph used in these experiments had a resolution of 2.5 cm^{-1} in the centre of the spectrum, and this increases towards the ends of the spectrum. It can be expected that more accurate results could be obtained using a spectrograph with better resolution.

A final point of note is the limited range over which the peaks in the polymer spec-

trum are distributed. The peaks that can be used in the wavenumber calibration protocol are distributed over a range from 743 cm^{-1}-1450 cm^{-1} while 4-Acetamidophenol spectrum contains peaks over a range 400 cm^{-1}-1650 cm^{-1}. We have found that the fit of the polynomial that is returned by the calibration protocol is less accurate outside of the left most and right most available peaks in the reference spectrum. Therefore, it may be expected that the wavenumber axis that is calibrated using the polymer reference would only be accurate within the range just mentioned. This may place a limit on the applicability of this material unless an improved protocol can be developed.

In this chapter, we have investigated a new wavenumber reference standard material for wavenumber calibration in the form of a photostable polymer. Chapter 5 will once again make use of the polymer investigated in this chapter as a wavenumber reference material, this time as part of new wavenumber protocol that will be shown to be superior to the third order fitting that has been used in this chapter. Chapter 6 will also make use of the polymer spectrum, but this time as part of a novel intensity calibration protocol. Before proceeding to these two chapters that will further investigate the polymer spectrum, in the next chapter, we take a step back and investigate a novel wavelength calibration protocol that makes use of an atomic emission spectrum. Raman wavenumber calibration can be achieved by first using wavelength calibration and then making use of the wavenumber conversion formula as described in Section 2.5; however, this requires an accurate knowledge of the laser wavelength.

Chapter 4

Improved wavelength calibration by modelling the spectrometer

The work in this chapter has been published in the following reference: *Liu, Dongyue, and Bryan M. Hennelly. "Improved Wavelength Calibration by Modeling the Spectrometer." Applied Spectroscopy (2022): 00037028221111796.* with the following abstract:

"Wavelength calibration is a necessary first step for a range of applications in spectroscopy. The relationship between wavelength and pixel position on the array detector is approximately governed by a low order polynomial and traditional wavelength calibration involves first-, second-, and third order polynomial fitting to the pixel positions of spectral lines from a well known reference lamp such as neon. However, these methods lose accuracy for bands outside of the outermost spectral line in the reference spectrum. We propose a fast and robust wavelength calibration routine based on modelling the optical system that is the spectrometer. For spectral bands within the range of spectral lines of the lamp we report similar accuracy to second- and third-order fitting. For bands that lie outside of the range of spectral lines we report an accuracy 12-121 times greater than that of third-order fitting and 2.5-6 times more accurate than second-order fitting. The algorithm is developed for both reflection and transmission spectrometers and tested for both cases. Compared with similar algorithms in the literature that use the physical model of the spectrometer, we search over more physical parameters in shorter time, and

obtain superior accuracy. A secondary contribution in this paper is the introduction of new evaluation methods for wavelength accuracy that are superior to traditional evaluation."

4.1 Introduction

In the previous chapter, we examined a new wavenumber reference material for direct wavenumber calibration in Raman spectroscopy. In this chapter, we change focus and look at wavelength calibration. Wavelength calibration is an important first step for various applications including astronomy, [118] multi-spectral imaging, [119, 120] and optical coherence tomography. [121, 122] Another application, Near-Infrared Spectroscopy (NIRS), which has widespread application in the identification of chemicals and biological materials, [123–125] requires wavelength calibration in order to produce reliable classification of spectra. [126–128]

Of particular importance in the context of wavelength calibration is Raman spectroscopy. Like NIRS, Raman spectra are commonly used to identify and classify materials based on large datasets of known spectra. Applications include pharmaceutical manufacture and bioprocess monitoring, [129, 130] material science, [131] and applications in clinical biology. [36, 132] Raman spectra have significantly higher resolution than NIRS spectra and, therefore, spectra must be subject to careful wavenumber and intensity calibration before comparison with a database. In Chapter 3, we looked at Raman wavenumber calibration using a Raman reference spectrum such as 4-acetamidophenol or a commercial polymer. However, as discussed in detail in Section 2.5 Chapter 2, which reviews wavenumber calibration in some detail, there are two classes of wavenumber calibration methods for Raman spectroscopy. One is to use direct wavenumber calibration with a reference material as in Chapter 3 and the other is to perform wavelength calibration followed by wavenumber conversion making use of the laser wavelength. As well as being a common precursor for wavenumber calibration, wavelength calibration is also commonly a first step intensity calibration [39, 59, 60], which is the core subject of Chapter 6 in this thesis.

Typically wavelength calibration involves polynomial fitting of the two dimensional dataset that is the known reference lamp spectral lines (wavelengths) and the position that these are found on the detector (pixels). [59, 60, 62, 63, 133] However, these methods tend to suffer from high error for regions outside of the spectral lines in the reference lamp, and this problem may be exacerbated for spectral bands for which there are few spectral lines available. Recently, there has been interest in using a physical model of the optical path in the spectrometer for the purpose of wavelength calibration, [78–82] which overcomes this limitation. All of these methods use the grating equation as the basis for developing an equation that relates the wavelengths and pixels in terms of the system parameters such as the grating period, spectrograph deviation angle, grating angle, camera pixel size and tilt. Some methods develop a system of simultaneous equations based on a set of wavelength, pixel pairs, and some are based purely on a brute force search over the various parameters in order to find the best fit of the equation to an available set of wavelength, pixel pairs. A thorough review of the existing state-of-the-art for wavelength calibration using both polynomial fitting and modelling the spectrometer, has been provided in Section 2.4 in Chapter 2.

In this chapter, we propose an algorithm based on the physical model that includes a brute force search for some of the system parameters, while performing polynomial fitting within that search to account for others. In doing so, we significantly reduce the scope of the search and improve the overall accuracy of the method. The reported accuracy is better than previous papers in this area. In addition, we provide several new evaluation methods that go much further than any previous publication in the area of wavelength calibration and we rigorously compare performance against polynomial fitting methods over large datasets. A more detailed list of the specific contributions in this chapter is provided in the next section.

4.2 Contributions in this chapter

In this chapter, a wavelength calibration method is proposed that is similar in design to that in Ref [79], with several differences:

- Similar to the method in Ref [79], our algorithm also searches over the variable parameters:
  1. The grating angle, which can often be electronically controlled.
  2. The centre position of the camera with respect to the optical axis.

  However, the physical model presented here also accounts for several more system parameters, including small errors in:
  3. The diffraction grating period and/or dispersion caused by displacement of the input irradiance spot vertically along the spectrometer slit. Image curvature is common in off-axis spectrometers [134] and leads to a deviation in the effective grating period. [81, 82]
  4. The angle of the optical axis with respect to a flat grating position.
  5. The focal length of the spectrometer .
  6. The camera pixel size
  7. A rotation of the camera plane.

  We note that the latter item relates to in plane rotation of the camera. Rotation of the detector plane with respect to the optical axis cannot easily be accounted for in the proposed algorithm, and care must be taken, experimentally, in order to ensure that slight defocusing of the spectrum irradiance does not occur on the detector. [88, 135]
- Although seven parameters are listed above, the algorithm proposed in this chapter does not employ a brute-force search over all of these, which would be intractable. Instead, a brute-force search is applied over a limited range of values for parameters (1), (3), and (4), only. The remaining parameters are all estimated using a simple ordinary least-squares fitting of a first-order polynomial within the three-dimensional brute-force search. The first-order is good enough to fix the possible existing small shift errors. The relationship between the parameters in

question is given by a simple straight line with unknown slope and zero-crossing. Therefore a simple linear regression applied to the data is sufficient to model this relationship. This is facilitated because parameters (2), (5), (6), and (7) will participate only in a simple shifting and scaling of the spectrum recorded by the detector. Therefore, the brute-force search space in our algorithm has one dimension less than that defined in Ref [79], while effectively searching over several more dimensions.

- The physical model is extended to account for spectrometers using both:
    1. A reflection grating. In this case, the physical model is based on a Czerny-Turner architecture. The resulting algorithm is tested on an Andor spectrometer with a rotating grating.
    2. A transmission grating. In this case the model is adapted for a Kaiser spectrometer with fixed volume holographic phase grating.
- The performance of the wavelength calibration algorithm is thoroughly investigated across a variety of gratings with different periods.
- The algorithm is rigorously evaluated using several different methods including 'leave-one-out' and 'leave-half-out', which provide a more accurate assessment of the calibration when compared to traditional approaches, particularly in spectral regions between the peaks and outside of end-peaks in the reference spectrum. Similar cross-validation approaches are commonplace in the field of chemometrics [136, 137] but we believe this is the first time they have been applied in the context of wavelength calibration.
- Finally, and most significantly, we report that the proposed method is significantly more accurate than any calibration method that we have so far reviewed in the literature for similar spectrometers. We report a standard deviation of <0.002nm, which appears to be approximately independent of grating period and resolution. High accuracy is maintained outside of the end-peaks of the reference spectrum.

4.3 Relationship between wavelength and pixel-position in a spectrometer

4.3.1 Physical model for generalized spectrometer with rotating grating

In this section, an equation is derived that relates the wavelength of a point-source at the spectrometer slit, to the position of the image of this point on the array detector. This derivation will form the basis of the calibration algorithm that is later developed in the following sections. The derivation is general for both transmission and reflection gratings, and the calibration algorithm can, therefore, be applied to spectrometers that employ both types of gratings as demonstrated in the subsequent subsections.

The diffraction grating is the main component of the spectrometer. The grating equation describes the relationship between the grating structure, the incident angle, and the angle of the diffracted light:

$$n\lambda = d\,(\sin\theta_\lambda \pm \sin\theta_i) \tag{4.1}$$

where d is the grating period, θ_i represents the angle of the incident ray of wavelength λ with respect to the grating normal, θ_λ is the angle at which this ray is diffracted, and n is the diffraction order. The $\pm$ term in the grating equation is negative for a transmission grating and positive for a reflection grating. Curvature of the slit image in the detector plane is caused by the displacement of the irradiance spot vertically along the slit resulting in an oblique angle of the light incident on the grating, and can be accounted for by adapting the grating equation as follows:

$$\begin{aligned} n\lambda &= d\cos\gamma\,(\sin\theta_\lambda \pm \sin\theta_i) \\ n\lambda &= d'\,(\sin\theta_\lambda \pm \sin\theta_i) \end{aligned} \tag{4.2}$$

where γ is the vertical oblique angle subtended by the optical axis and the line connect-

ing the centre of the collimating lens (or mirror) and the vertical position of the spot on the slit, [81, 82, 134, 138] and $d' = d\cos\gamma$.

Spectrometers often employ a rotating grating such that different wavelength bands can be projected onto a fixed detector. In the case that the grating is rotated by an angle θ_d, both the incident and diffraction angles will be altered. This is illustrated in Fig. 4.1 in which a reflection grating is mounted on a rotating triangular base; this is similar to the design of one of the two spectrometers that is investigated later. The blue image represents the initial state of the spectrometer, without rotation, for which the incident ray is propagating at an angle α with respect to the grating normal. The zero-order diffracted ray (also at angle α with respect to the grating normal) propagates through the centre of a lens of focal length f, and on to the centre of a detector array; for simplicity, we will later refer to this as the optical axis of the spectrometer. We note that the value 2α is often referred to as the deviation angle of the spectrometer. We also note that the focusing optic can also take the form of a parabolic mirror as described in the following section. The black image represents the state of the spectrometer following rotation of the grating by an angle of θ_d. For the same incident ray we derive the position of the resulting n^{th} order diffracted ray on the detector.

The grating equation can be rewritten to describe diffraction by the rotated grating as follows:

$$\frac{n\lambda}{d} = \sin(\beta - \theta_d) + k\sin(-\alpha - \theta_d) \tag{4.3}$$

The parameter β in the figure represents the angle of the diffracted ray with respect to the grating normal for the initial state. The $\pm$ symbol has been replaced with the parameter k, which takes the value of $+1$ and -1 for transmission and reflection gratings, respectively. The angle β can be defined in terms of the other parameters as follows:

$$\beta = \sin^{-1}\left[\frac{n\lambda}{d'} - k\sin(-\alpha - \theta_d)\right] + \theta_d \tag{4.4}$$

The angle between the n^{th}-order diffracted ray and the optical axis is $\beta - \alpha$. The position

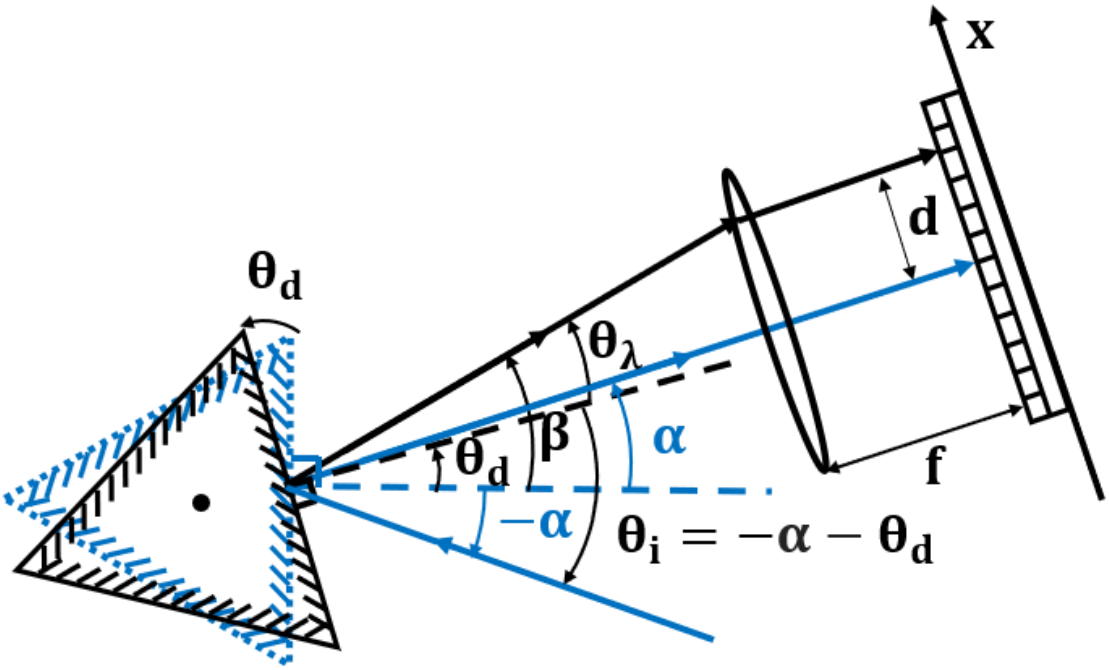


Figure 4.1: Diffraction of a ray by a rotated grating. The blue illustration shows the zero-order diffraction of an incident ray onto the optical axis of the spectrometer for a flat grating position. The black image shows the −1 order diffraction of the same ray following rotation of the grating by an angle θ_d. We stipulate that the counterclockwise direction is positive for all angles.

at which this ray will be incident on the detector array is given by:

$$xT = f\tan(\beta - \alpha) + C \tag{4.5}$$

where C represents misalignment of the centre of the detector array with respect to the optical axis, and T is the pixel pitch in the detector. Equation 4.5 can be rewritten as follows:

$$x = \frac{f}{T}\tan\left\{\theta_d + \sin^{-1}\left[\frac{n\lambda}{d'} - k\sin(-\alpha - \theta_d)\right] - \alpha\right\} + \frac{C}{T} \tag{4.6}$$

$$\lambda = \frac{d'}{n}\left\{\sin\left[\tan^{-1}\left(\frac{xT - C}{f}\right) + \alpha - \theta_d\right] + k\sin(-\alpha - \theta_d)\right\} \tag{4.7}$$

Equations 4.6 and 4.7 is the basis of the calibration algorithms that are proposed in later sections. Before these algorithms are described, we first explore the nature of the relationship between the wavelength, λ, and pixel-position, x, as defined by Equa-

Parameter	Unit	Czerny-Turner	Transmission
Reflection/ transmission (k)	NA	+1	-1
Diffraction order (n)	NA	-1	+1
Grating Period (d)	lines/mm	300 600 1000	2455
Half the deviation angle (α)	Degree	10.94	45
Grating angle (θ_d)	Degree	4.8 10.2 17.2	0
Focal length (f)	mm	500	85
Camera pixel pitch (T)	µm	26	26
Camera width (N)	pixels	1024	1024
Camera centre position (C)	pixels	0	0

Table 4.1: The parameters for the two spectrometers illustrated in Fig. 4.2, which are investigated in this study. The parameters correspond to those in Equation 4.6.

tion 4.6, for two spectrometers, which are later the subject of the proposed calibration algorithms. The first spectrometer of interest is a Czerny-Turner spectrometer employing parabolic mirrors and three different plane-ruled reflection gratings. The second is a lens based spectrometer employing a volume-phase holographic transmission grating. Both spectrometers are described in more detail below, followed by a discussion on the application of Equation 4.6 to model each system.

4.3.2 Reflection spectrometer

A traditional Czerny-Turner spectrometer with focal length 500mm and with a motorized rotating grating was utilised for most of the experiments reported in this chapter (Andor Shamrock 500; SR-500i-A; Andor UK), which is illustrated in Figure 4.2 (a).

Converging light enters the spectrometer slit and is collimated by a parabolic mirror and directed onto a grating, housed on triple grating turret and mounted on a rotation stage. The three gratings on the turret are all plane-ruled reflection gratings with the

following specifications: 1000 lines/mm with blaze at 900nm (Andor SR5-GRT-1000-0900; Andor UK), 600 lines/mm with blaze at 750nm (Andor SR5-GRT-0600-0750; Andor UK), and 300 lines/mm with blaze at 760nm (Andor SR5-GRT-0300-0760; Andor UK). The angled grating directs the $n = -1$ diffraction order towards a second parabolic mirror, which focuses the image of slit at the detector plane. The detector is a cooled CCD (Andor iDus; DU420A-BR-DD; Andor UK) with 256×1024 pixels with a pixel-pitch, T of 26μm. Both parabolic mirrors have a focal length, f of 500mm and the half deviation angle, α, was measured to be 21.88°. The values of each parameter in Equation 4.6 for this spectrometer are provided in Table 4.1; grating angles, θ_d are selected for each grating.

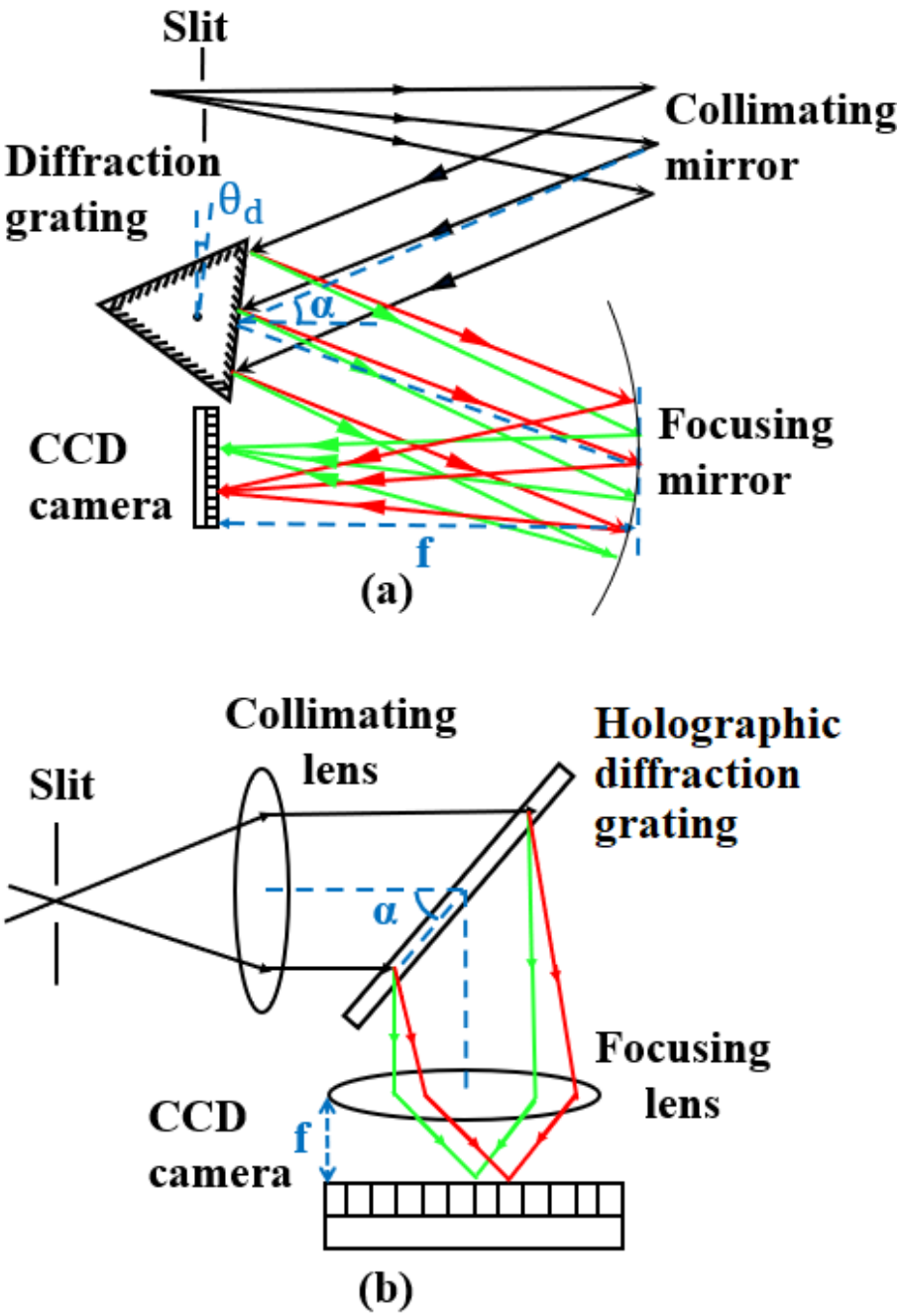


Figure 4.2: (a)The Czerny-Tuner spectrometer using parabolic mirrors and a rotating grating, and (b) A transmission spectrometer utilising glass lens focussing and a holographic grating. A Holographic diffraction grating is used for holographic grating. The proposed wavelength calibration algorithm is general such that it can be applied to both types of spectrometers.

4.3.3 Transmission spectrometer

A transmission spectrometer (HOLOSPEC-F/1.8I-VIS; Andor, UK) is also investigated in this study, the design of which is illustrated in Figure 4.2 (b). Light is input to a slit of width 25μm. A first lens collimates the light and is followed by a volume-phase

holographic transmission grating with 2455 lines/mm (HS-HSG-532-LF; Andor, UK), which is angled at $\alpha = 45°$ with respect to the optical axis. A second lens captures the $n = +1$ diffraction order and images the slit onto the detector, a cooled CCD (Newton DU920P-BVF; Andor, UK) with 256×1024 pixels and pixel-pitch, T, 26μm. Both lenses have a focal length, f, of 85mm. Notably, in this case, the diffraction grating is fixed and the grating angle is $\theta_f = 0°$.

4.3.4 Relationship between wavelength, λ, and pixel-position, x, for both spectrometers

In the sections that follow, a wavelength calibration algorithm is proposed that exploits the relationship between the wavelength, λ, and pixel-position, x, for a given spectrometer, based on the model described above in Equation 4.6. Here, we first explore the nature of this relationship for the reflection and transmission spectrometers that are described in the proceeding subsections

This relationship is shown in Fig. 4.3, where the values in Table 4.1 are substituted into Equation 4.6 for all four gratings. For ease of comparison, the wavelength axis has been normalised such that the minimum and maximum values appearing on the extreme ends of the CCD are 0 and 1 for all four diffraction gratings. A dashed line shows a linear relationship between x and λ. Interestingly, the 600 lines/mm grating is shown to exhibit the most linear relationship between x and λ, while the 300 and 1000 lines/mm gratings are both less linear but appear on opposite sides of the straight line. The 2455 lines/mm grating is significantly non-linear primarily owing to the shorter focal length and higher dispersion. The variability in the linearity of the relationship as a function of grating period and focal length may explain why several different polynomial orders have previously been proposed to be optimal for fitting as a means of wavelength calibration by different authors as discussed in the introductory section above.

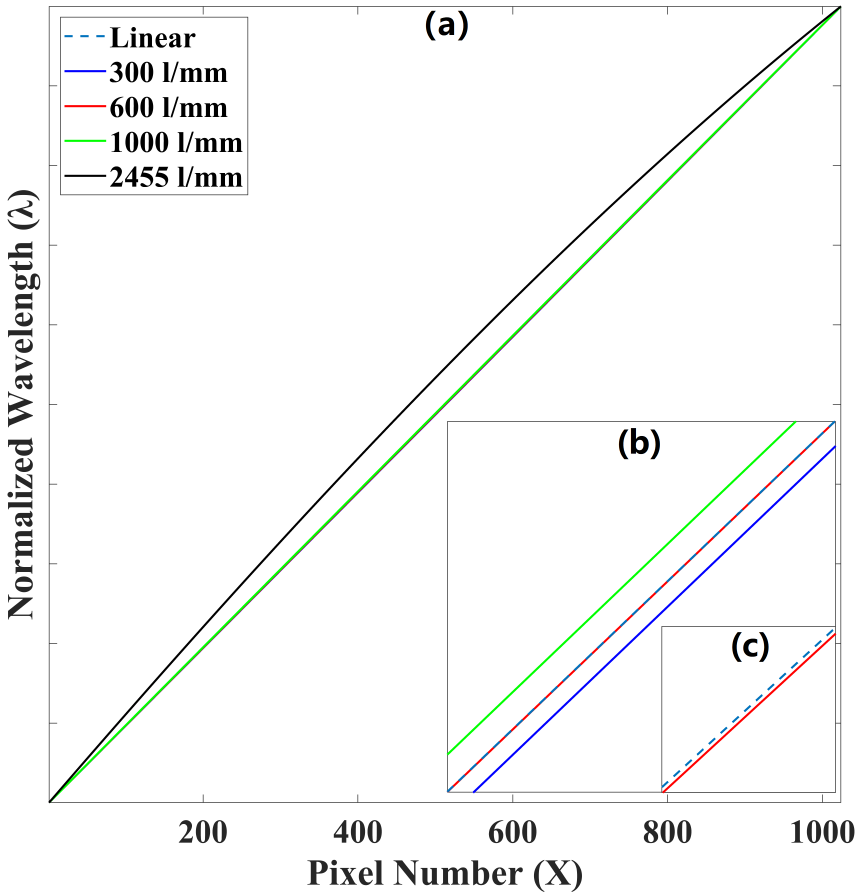


Figure 4.3: Investigation of the non-linearity of the (x,λ) relationship for the four different gratings that are later used for testing. (a), (b), and (c) show increasingly zoomed in areas. These plots are based on Equation 4.6 using the parameters listed in Table 2. The wavelength axis has been normalised for direct comparison.

4.4 Calibration based on the physical model

Here we describe the sequence of steps that comprise the proposed wavelength calibration algorithm which is general to either the transmission or reflection spectrometer. We begin by clearly posing the problem and this is followed by describing two algorithms that can be used to solve this problem. The parameter d' is equal to $d\cos(angle)$ where d is the grating number and the angle is with respect to the horizontal. Both can be coupled together in the search. S is the set of parameters that define the system as described earlier:

$$S = [f, T, C, d', \alpha, \theta] \tag{4.8}$$

The relationship between the pixel coordinates on the detector, and the corresponding wavelength values that these pixels capture, is predicted by the physical model defined in Equation 4.6 in terms of the parameter set S. This equation, and its inverse given by Equation 4.7, which relates wavelength to pixel, are summarised by the following equations:

$$\begin{aligned} x &= model(S, \lambda) \\ \lambda &= model^{-1}(S, x) \end{aligned} \tag{4.9}$$

We move now from a continuous model to a discrete one, where the values of x and λ belong to the two sets of discrete values defined as follows:

$$\begin{aligned} X_0 &= [x_1^0, x_2^0, \cdots, x_N^0] \\ \lambda_0 &= [\lambda_1^0, \lambda_2^0, \cdots, \lambda_N^0] \end{aligned} \tag{4.10}$$

where λ_0 is a set of known neon peak wavelengths of which there are N, and X_0 is the corresponding set of positions in the detector plane at which these peaks are detected; we note that the values of X_0 will not in general be integers; a pre-processing step is first implemented in order to estimate the sub-pixel position of a peak as described in the following section.

The goal is, therefore, as follows. We wish to design an algorithm that can determine the set of parameters, S, that 'best' relate the known N wavelength values, λ_0, and

corresponding pixel positions at which these wavelengths were measured, X_0, according to Equation 4.9. We begin by defining a brute-force algorithm based on Equation 4.9 that is conceptually simple but computationally intractable. This algorithm is used as the basis of a second algorithm, which is significantly more computationally efficient.

4.4.1 Algorithm 1: Brute force

The first algorithm is based on a simple but computationally expensive brute-force search overall of the parameters in S and is made up of three steps:

1. The first step provides initial estimates of the key parameters in S, which are defined as $S_0 = [f_0, T_0, C_0, d_0, \alpha_0, \theta_0]$ as follows:
   - The values of (f_0, T_0, d_0) can be taken from the manufacturers specifications for the spectrograph and detector, where we assume $d_0 = d$ and $\gamma = 0$.
   - α_0 is measured manually.
   - C_0 and θ_0 are estimated using a brute-force search over only these two variables for a single peak pair (x_i^0, λ_i^0).
2. The second step is to perform a brute-force search over all six parameters in S over some range/step-size centred at S_0. Each unique set of parameters, S_j, in this range will produce set of pixel positions X_j as follows:

$$X_j = model(S_j, \lambda_0) \tag{4.11}$$

   The specific set of parameters, S_{min}, that produces the set X_{min} that most closely match the actual pixel values X_0 at which the peaks are detected are taken to be the true system parameters. This is determined by minimising the error function defined in Equation 4.12 over all parameter sets j in the range of the search.

$$err = \sum_{i=1}^{N} (x_i^j - x_i^0)^2 \tag{4.12}$$

   We acknowledge that a tight-grid brute force search over such an error metric would not normally be applied since more efficient search algorithms are far more efficient such as steepest descents and simplex searching. Algorithm 1 serves

only as a natural introduction to Algorithm 2, which must employ a brute force search albeit over a much smaller range of values, and for this reason it is defined in terms of a brute force search algorithm.

3. Now that the system parameters S_{min} have been found, these can be used to relate the integer pixel (centre) positions to the corresponding wavelength values, thereby providing wavelength calibration for the spectrograph. This third and final step is defined in the equation below:

$$\lambda_{cal} = model^{-1}[S_{min}, [1,2,\cdots,1024]] \tag{4.13}$$

where λ_{cal} represents the set of calibrated wavelength values associated with each pixel centre position. [1,2,$\cdots$,1024] denotes the integer set of pixels.

While this algorithm provides for accurate calibration, it requires a brute-force search over six parameters and is computationally intractable. Even making the somewhat reasonable assumption that the specifications for $d' = d$, and T, the camera pixel size, are without any error, will require a four-dimensional search, which remains time-consuming.

4.4.2 Algorithm 2: Speed-Up Using Least-Squares

In this section we attempt to speed-up the running time of *Algorithm 1* by using the classical least-squares algorithm. Referring to Equation 4.7, it is clear that the parameters f, T, C perform only scaling and additive functions on the spatial coordinate x and can, therefore, be accounted for using linear regression. Therefore, a brute-force search is required only over the remaining parameters α, d, and θ. The second algorithm also contains three steps as follows:

1. This is identical to Step 1 in Algorithm 1.
2. Here a brute-force search is performed over only a three parameter set, $[\alpha, d', \theta]$, over some range of values, centred at $[\alpha_0, d_0, \theta_0]$ and using the values f_0, T_0, C_0 in

order to provide an intermediate result.

$$X_j = model([\alpha_j, d_j, \theta_j, C_o, f_o, T_o], \lambda) \tag{4.14}$$

For each unique set of values $[\alpha_j, d_j, \theta_j]$ over the search range, the resultant values X_j are linear-regressed with respect to positions at which the peaks were detected, X_0, in order to account for errors in f, T, and C, which provides an updated set of estimated positions X_j. For simplicity, we describe this operation in terms of the Matlab functions, `polyfit` and `polyval`, which are used to implement it as follows:

$$\begin{aligned} P_j &= polyfit(X_j, X_0, n) \\ X_j &= polyval(X_j, P_j) \end{aligned} \tag{4.15}$$

where the function `polyfit` returns the coefficients of degree n that is the best fit (in a least-square sense) to describe the transformation between X_j and X_0. This is followed by the function `polyval`, which applies this transformation to X_j using these coefficients in order to provide the updated values for X_j.

The specific set of parameters, $S_{min} = [\alpha_{min}, d_{min}, \theta_{min}, f_0, T_0, C_0]$, and linear regression defined by P_{min} that produces the set X_{min} that most closely match the actual pixel values X_0 at which the peaks are identified and are taken to be system parameters. Once again, this is determined by minimising the error function defined in Equation 4.12 over all parameter sets j in the range of the search.

3. Now that the system parameters S_{min} have been found, as well as the coefficient for the linear regression that accounts for error in f, C, and T, the integer pixel (centre) positions can be related to the corresponding wavelength values, thereby providing wavelength calibration for the spectrograph. In this final step, the pixel position are projected into the wavelength domain by using the opposite process outline in Step 2:

$$\lambda' = model^{-1}[S, polyval([1, 2, \cdots, 1024], P_{min})] \tag{4.16}$$

Here the camera pixels are defined in terms of integers $1 \rightarrow 1024$.

4.5 Overall calibration procedure

As discussed in the previous sections, the core principle of wavelength calibration of a spectrometer is to first record a reference spectrum containing some number of sharp, symmetrical and well defined, known peak wavelengths. The second step is to identify the pixel positions of the various peaks in the recorded reference spectrum, which can then be used in the third step, which involves fitting with either a low-order polynomial or the pixel-wavelength relationship defined by a physical model. In either case, a matching wavelength value must be assigned to each pixel in the detector. The minimum number of requisite peaks in the recorded reference spectrum depends on the fitting method; a first-order polynomial fitting requires only two peaks, with this number increasing with respect to the polynomial order used. For the two physical model based methods reviewed earlier, there is also a minimum number of peaks required; for example the brute-force method in Ref. 79 can work with a minimum of four peaks, while the method based on simultaneous equations [80] requires a minimum of five peaks. A simple rule of thumb is that there must be at least as many peaks in the reference as there are variables in the physical model or coefficients in the polynomial. In general, however, more accurate results are obtained by increasing the number of peaks in the reference. As well as requiring a large number of peaks, the distribution of these peaks must also be considered. As noted by previous authors, [59], wavelength calibration using a polynomial order greater than one, will result in poor calibration for bands that lie outside the end peaks at either side of the reference spectrum. This is because there are no peaks in these extreme regions that can constrain the polynomial coefficients. However, first-order polynomial fitting (which is rarely accurate to begin with), and fitting based on the physical model are more robust to these out-of-band calibration errors.

Typical reference lamps that are used for wavelength calibration include, mercury-argon, neon, krypton, which are often selected based on the number of peaks available

in the band of interest. The latter two reference lamps are utilised in this chapter. In Fig. 4.4 (a) the spectrum of the neon lamp (Spectrum tube-neon gas; Edmund Optics UK) is shown, recorded using the Czerny-Turner spectrometer described earlier using a 300 lines/mm grating. Also shown in the figure are the bands of peaks that can be captured by the 600 lines/mm, and 1000 lines/mm grating, which can be moved by rotation of the grating angle. In Fig. 4.4 (b) the spectrum of the krypton lamp (Spectrum tube-krypton; Edmund Optics UK) is shown, recorded using the transmission spectrometer described earlier using a 2615.8 lines/mm grating. In this case the grating angle is fixed and the spectrometer can record only the band 530-610 nm. The bandwidth of this spectrometer and the Czerny-Turner spectrometer with the 1000 lines/mm grating are similar due to the significantly different focal lengths in the two spectrometers. It is notable that wavelength calibration of this spectrometer with the krypton map with polynomial fitting with order two or more will result in significant error in the left-most band 530-556 nm due to the absences of peaks in this band. This 'error band' would increase further using the neon lamp since the first useful peak occurs at 585 nm. In Table 4.2 the exact peak wavelengths for these two sources are shown, which have been taken from the database of the National Institute of Standards and Technology. [1]

In order to achieve accurate calibration, identification of the peak position requires sub-pixel resolution even though such a resolution is in general less than the specified resolution of the spectrometer. Various methods have been proposed in the literature to achieve such accuracy, including upsampling of the reference spectrum by zero-padding the discrete Fourier transform of the spectrum [39, 62] as well as fitting a Lorentzian function, or similar, to the pixel values in the region of the peak. [43, 71, 77] We have tested these various approaches and determined that fitting with a Lorentzian function[139] of the following form is slightly more accurate than upsampling:

$$Peak' = \frac{P_1}{(x_{range} - P_2)^2 + P_3} + P_4 \tag{4.17}$$

where Peak means P_1, P_2, P_3 and P_4 are fit-parameters, x_{range} represents the pixel range of the peak, $Peak'$ represents the new intensity value in this pixel range.

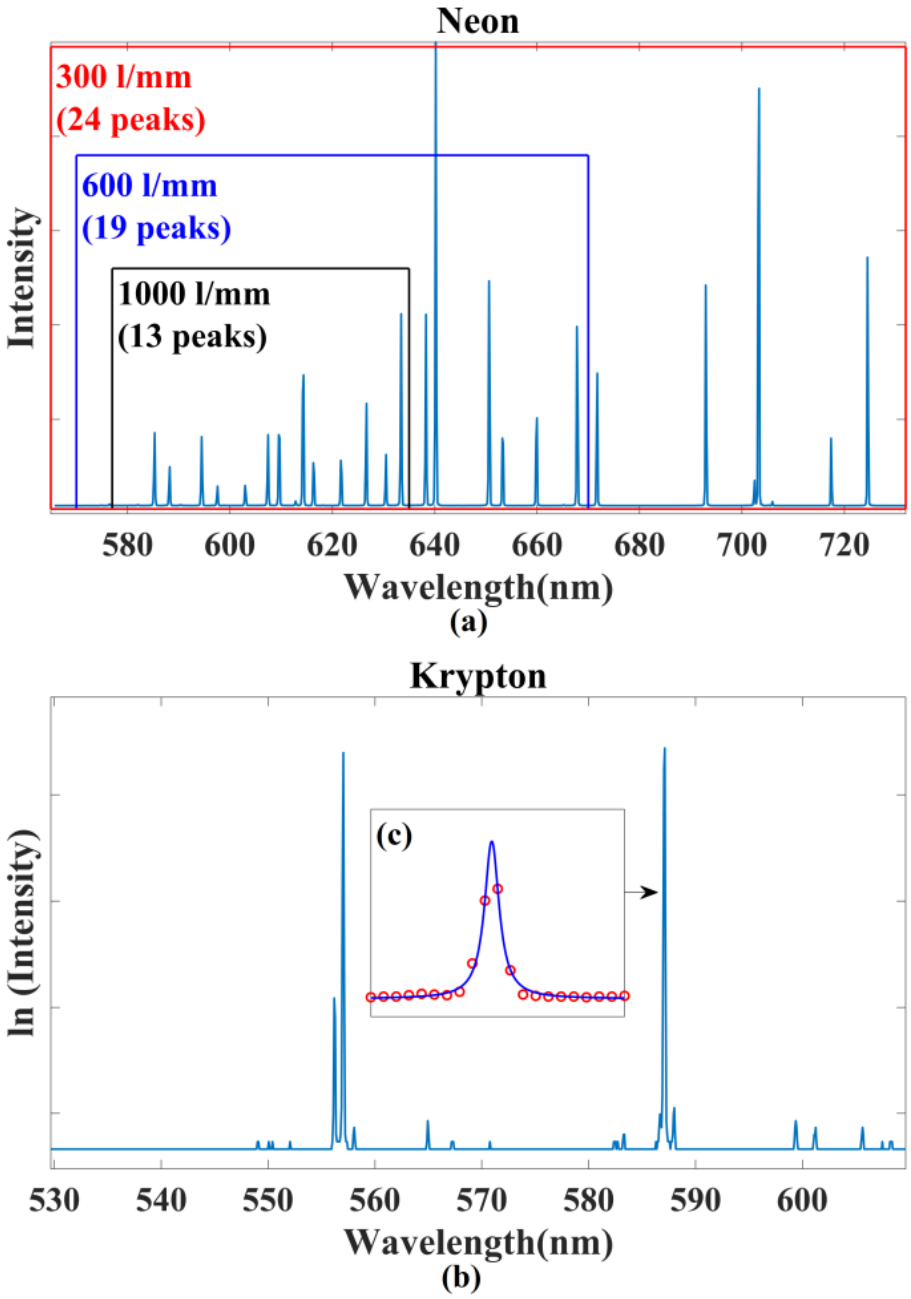


Figure 4.4: The spectrum of (a) neon (captured by the Czerrny-Turner spectrometer) and (b) krypton (captured by the transmission spectrometer). (c) A single krypton peak is shown illustrating the method of peak fitting for sub-pixel accuracy.

An example of this approach is shown in Fig. 4.4 (c) in which we show a Lorentz function that has been fit to one peak in the krypton spectrum. In the results section below, all of the the peak positions in each reference spectrum are estimated with sub-pixel accuracy using this approach.

The overall procedure can be divided into four steps. The first step is to record the reference spectrum, and the second step is to identify the sub-pixel position of

Wavelength / nm		
Neon		
585.2488±5*e*-5	588.1895±5*e*-5	594.4834±5*e*-5
597.5534±5*e*-5	602.9997±5*e*-5	607.4338±5*e*-5
609.6163±5*e*-5	614.3063±5*e*-5	616.3594±5*e*-5
621.7281±5*e*-5	626.6495±5*e*-5	650.6528±5*e*-5
630.4789±5*e*-5	633.4428±5*e*-5	638.2991±5*e*-5
640.2248±1*e*-4	653.2882±5*e*-5	659.8953±5*e*-5
667.8277±5*e*-5	671.7043±5*e*-5	692.9467±4*e*-5
703.2413±4*e*-5	717.3938±4*e*-5	724.5167±4*e*-5
Krypton		
556.2225±4*e*-5	557.0289±4*e*-5	558.0387±4*e*-5
564.9562±5*e*-5	567.2451±5*e*-5	570.7513±5*e*-5
583.2857±5*e*-5	587.0916±5*e*-5	587.99 ±5*e*-5
599.385 ±5*e*-5	601.2156±5*e*-5	605.6126±5*e*-5

Table 4.2: Reference spectral lines used in this chapter (with uncertainties [1])

each spectral peak that is listed in the related reference database as described above. The third step is the application of Algorithm 2 described earlier, which returns the parameters for a single equation that relates wavelength to pixel position. The final step is to apply this equation in order to identify the wavelength associated with the centre of each pixel. For traditional calibration, the third step would be replaced with polynomial fitting to find the coefficients of an n-order polynomial and the fourth step would be application of this polynomial to identify the wavelength for each pixel.

4.6 Experiment design

4.6.1 Recording of reference spectra

In total, we examine the performance of the proposed algorithm across two spectrometer designs and four different gratings periods as described earlier, each with varying dispersion. For the case of the Czerny-Turner system, only the reference neon lamp is applied and for the case of the transmission lens spectrometer, only the krypton lamp is applied; typical spectra are shown in Fig. 4.4 (a) and (b) for both cases and in (c) an example of fitting the Lorentzian peak is shown; this achieves sub-pixel accuracy as described in the previous section. The Czerny-Turner spectrometer is investigated using a 300 lines/mm, 600 lines/mm, and 1000 lines/mm corresponding to different wavelength bands as illustrated in Fig. 4.4 (a), while the transmission spectrometer uses a grating with 2455 lines/mm. For each of the three gratings in the Czerny-Turner spectrometer, 100 different reference spectra are recorded with slight movements of the grating rotation angle. For these three cases, a rigorous evaluation of the performance of the calibration is possible by calculating the ensemble average of the error metrics defined below, across the set of 100 reference spectra.

For all cases, the lamp was first carefully centred on the slit to ensure symmetrical spectral peaks. Andor Solis software is used to record the raw spectra in the image plane. Because of the strong irradiance from the lamps, a diffuser was positioned between the lamp and slit. To reduce the effect of noise, the accumulation time was varied to provide a photon count that was just less than the saturation level of the CCD. Rather than use Full Vertical Binning, which can produce error in the presence of image distortion, images were recorded from the detector as shown in Fig. 4.5 for both the (a) neon and (b) krypton lamps. The centre row of pixels was cropped as illustrated by the red box in the figures. This approach was taken instead of Full Vertical Binning, in order to overcome the problem of image distortion as described earlier. With vertical Vertical binning, pairs of adjacent pixels from two lines in the sensor are summed. As the read-out noise of each read operation is now just applied to less resulting pixel information

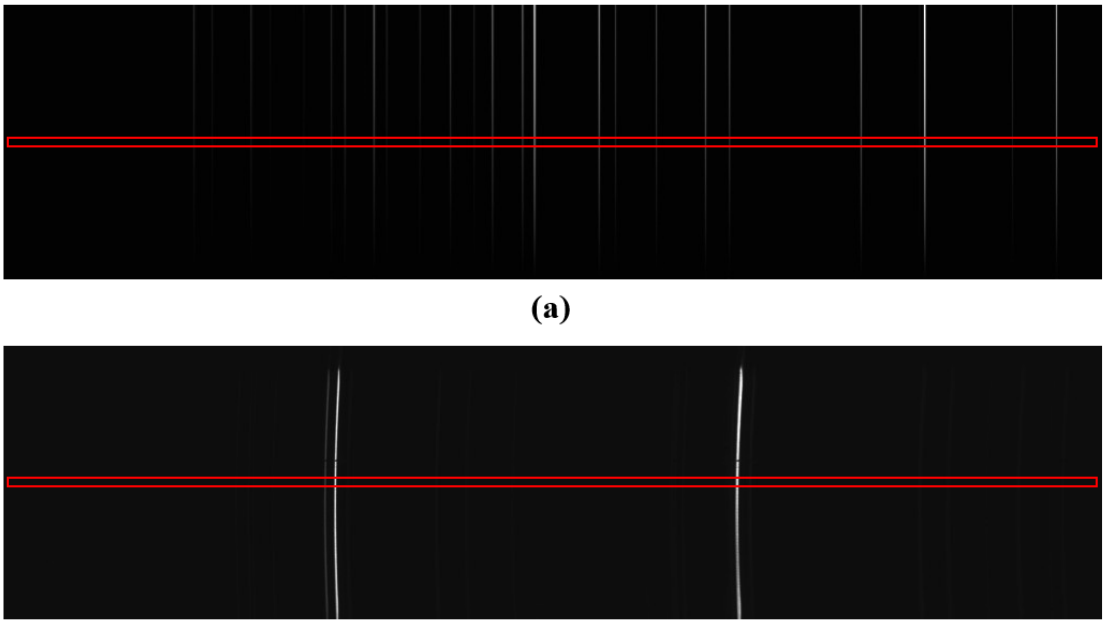


Figure 4.5: Imaging the reference spectra in the detector plane: (a) neon-300 lines/mm (b) krypton-2455 lines/mm; for the latter case clear distortion is observed due to the effect of the lens. A cropped row of pixels is extracted to mitigate this effect.

the signal-to-noise ratio (SNR) is increased because of the reduced noise at the higher combined signal.

4.6.2 Error metrics

For comparison with similar methods proposed in the literature, several different error metrics are reported including, mean absolute error (MAE), the standard deviation (SD), and the root mean square error (RMSE), all of which have appeared in different papers. These three metrics are defined below. A calibrated reference peak wavelength is denoted as (λ_i), and assuming N such reference peaks exist in the reference spectrum, the following error metrics are defined:

$$error(\lambda_i) = calibrated(\lambda_i) - NIST(\lambda_i) \tag{4.18}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |error(\lambda_i)| \tag{4.19}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} |error(\lambda_i)|^2} \tag{4.20}$$

$$SD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} |error(\lambda_i) - ME|^2} \tag{4.21}$$

In order to provide a more reliable estimate of the error, the above metrics are calculated for a set of M different spectra where the grating is moved between captures. The ensemble average of each of the above metrics is calculated over these M reference spectra as follows:

$$\overline{MAE} = \frac{1}{M}\sum_{k=1}^{M} MAE(k) \tag{4.22}$$

$$\overline{RMSE} = \frac{1}{M}\sum_{k=1}^{M} RMSE(k) \tag{4.23}$$

$$\overline{SD} = \frac{1}{M}\sum_{k=1}^{M} SD(k) \tag{4.24}$$

4.6.3 Evaluation methods

We employ three methods of evaluation that employ the metrics listed above, two of which are proposed for the first time.

1. *All-peaks*: Here all of the calibrated peaks from the reference are used in the error analysis. This is by far the most common approach in the literature.
2. *Leave-one-out-cross-validation*: in order to remove any bias from the reference spectrum, we propose for the first time in the field of wavelength calibration (to the best of our knowledge) the use of cross-validation, an approach that is borrowed from the field of chemometrics. [136, 137] For the first case, 'leave-one-out' cross-validation, one peak is removed from the reference spectrum used in the calibration process. The error metric is then applied only to this peak after calibration. This process is repeated for each peak in the spectrum and the average value for all cases is calculated. We believe that this is the first time that such an approach has been taken and we expect that it will provide a more accurate estimate of wavelength accuracy within the band of spectral lines provided by the reference lamp.
3. *Leave-half-out*: Similar to the approach taken in Ref [79] we propose an evaluation based on calibrating using the left-most half of the reference peaks and apply-

ing the error metric to the right-most peaks of the calibrated spectrum. This is repeated using the right-most peaks for calibration and the left-most for error calculation. The average of the two values is taken. The advantage of this approach is that the accuracy of calibration is tested in bands outside of the outermost end peaks in the reference lamp; the other two methods of evaluation only test for accuracy within the bounds of the reference spectrum lines.

4.6.4 Comparison with traditional methods of wavelength calibration

In all cases, the proposed algorithm is compared with equivalent results from first-order, second-order, and third-order polynomial fitting and several interesting conclusion are made in the following section concerning the accuracy of these different methods under different conditions. Fourth-order fitting and higher provided no improvement in results and is not presented here.

4.7 Results

In this section, the results are presented for wavelength calibration using Algorithm 2 and compared with the corresponding set of results from first-, second, and third-order polynomial fitting. These results are broken down into three sets of evaluations, corresponding to 'All-Peaks' (ALL), 'Leave-one-out-cross-validation' (LOO), and 'Leave-half-out' (LHO). Furthermore, to facilitate comparison with other papers, which use various metrics, these evaluations are performed using three different metrics: $\overline{MAE}$, $\overline{RMSE}$, $\overline{SD}$ as defined in Equations 5.18, 5.19, and 5.20. For the case of the transmission spectrometer, the grating angle could not be adjusted and so only a single spectrum was available. In this case the error metrics used in the evaluation are: MAE, $RMSE$, SD as defined in Equations 6.6, 5.16, and 5.17. The results for the Mean Absolute Error metrics are shown below in Fig. 4.6. It can be seen that the traditional evaluation method of inspecting all peaks provides approximately $10-20\%$ superior results compared with LOO, which is proposed for the first time in this chapter, and which we believe is a more accurate representation of wavelength calibration within the range of wavelength defined by the outermost reference lamp spectral lines. However, the overall trend of the results are the same for both ALL and LOO. It can be seen for both of these evaluation methods, that first-order fitting is the worst method in all cases but provides its best result for the 600 lines/mm grating, which was earlier shown to produce the most linear relationship between wavelength and pixel position (see Fig. 4.3). For the case of LOO evaluation, Algorithm 2 provides equivalent results to second- and third- order fitting for the 300, 600, and 1000 lines/mm gratings with very little difference between the three cases: (0.016 nm error for the 300 lines/mm case and 0.006 nm error for the other two). For the case of the 2544 lines/mm grating, which has by far the most non-linear relationship between wavelength and pixel position, third order fitting provides the best LOO accuracy with an error of 0.00498 nm, and Algorithm 2 provides the next best LOO accuracy with an error of 0.00897 nm. However, it should be noted that this case uses only a single spectrum and only 12 krypton peaks were available. More conclusive results could not be obtained by rotating the grating into different states as for the other

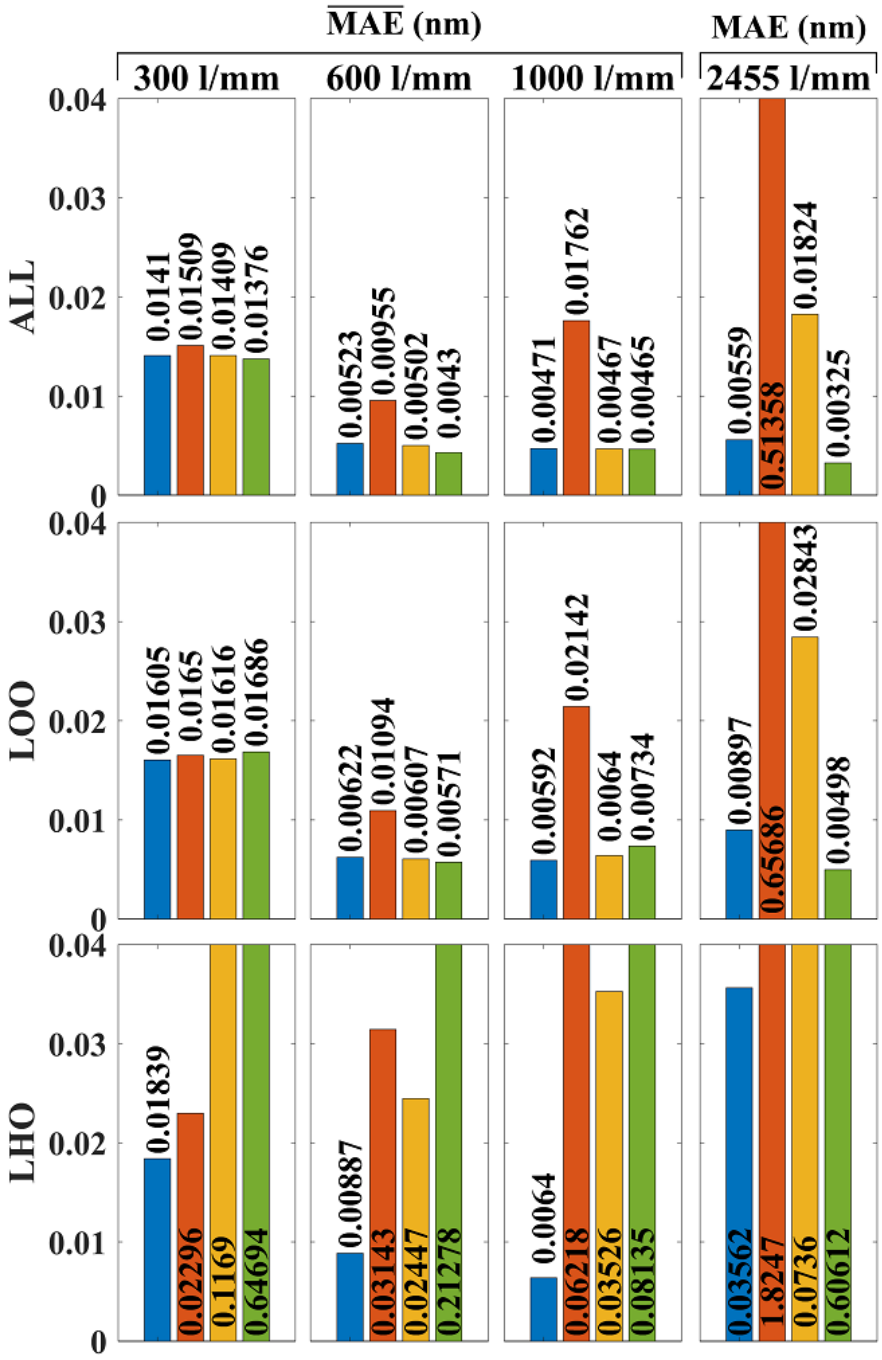


Figure 4.6: Evaluation of wavelength calibration accuracy using Mean Absolute Error. A neon reference lamp is used for the Crezny-Turner reflection spectrometer with three different gratings: 300, 600 1000 lines/mm and for these three cases the $\overline{MAE}$ error metric is applied over 100 spectra with grating movement between capture. A krypton reference lamp is used for the transmission spectrometer with grating 2455 lines/mm and for this case, the MAE error metric is applied over a single spectrum. The results of Algorithm 2, proposed in this chapter, is given in blue and the results for first-, second-, and third-order polynomial fitting are given in orange, yellow, and green, respectively. The results of 'All-Peaks' (ALL), 'Leave-one-out-cross-validation' (LOO), and 'Leave-half-out' (LHO) are shown on different rows. For ease of comparison, the same axis range is used for all three evaluations. In several cases, the bars have been capped at 0.04 nm to improve visualisation. The correct values are overlaid on the bars in all cases.

three gratings.

The superiority of Algorithm 2 is evident for the third evaluation method, LHO, which provides a more accurate estimate of error in regions that are outside of the bandwidth of the reference lamp spectral lines. For the case of the 300 lines/mm reflection grating, Algorithm 2 provides the best LHO accuracy, with an error of 0.01839 nm and first-order fitting is next best with an error of 0.02296 nm; second- and third-order fitting error are 6-times and 35-times worse than that of Algorithm 2, respectively. For the 600 lines/mm reflection grating, Algorithm 2 once again provides the best LHO accuracy with an error of 0.00887 nm; first-, second-, and third- order fitting errors are 2.5-times, 3.5-times, and 24-times greater than that of Algorithm 2, respectively. For the third reflection grating of period 1000 lines/mm, Algorithm 2 once again returns by far the best LHO accuracy with an error of 0.0064 nm; first-, second-, and third-order fitting provide errors that are 10-times, 4-times, and 12-times greater than that of Algorithm 2, respectively. Notably, when Algorithm 2 is used to calibrate the reflection spectrometer, leave-half-out evaluation provides similar results when compared with leave-one-out evaluation; there is only a marginal increase in error of 10-30% for the former, indicating that Algorithm 2 provides similar results far outside of the reference lamp spectral lines, as it does within the bandwidth of the lamp. This is not the case for the polynomial fitting; while third-order fitting provides equivalent results to Algorithm 2 for wavelengths within the bandwidth of the reference lamp (as evidenced by LOO evaluation), the error increases by a factor of 12-35 in regions outside of the lamp bandwidth (as evidenced by LHO evaluation).

For the case of the transmission grating with period 2455 lines/mm, all methods fare worse for LHO evaluation when compared with LOO evaluation; it can be seen that Algorithm 2, first-, second-, and third-order fitting provide LHO error that are 4-, 2.5-, 2.5-, 121-times greater than the corresponding LOO error. This is likely due to the small number of peaks available from the krypton lamp in the band of interest, which is exacerbated for LHO evaluation. Regardless, Algorithm 2 is the most accurate with an error of 0.03562 nm; first-, second-, and third-order fitting provide errors that are

51-times, 2-times, and 17-times greater than that of Algorithm 2, respectively.

It is important to record the accuracy of the calibration methods in the context of the spectrometer resolution. The Czerny-Turner spectrometer with 300, 600, and 1000 lines/m grating is specified to have a resolution of 0.32 nm, 0.15 nm, and 0.09 nm, respectively and the transmission spectrometer provides a resolution of 2.97 nm. All of these resolutions are significantly larger than the accuracy provided by Algorithm 2. Equivalent results are shown in the appendix using standard deviation and RMSE in place of the MAE metric.

4.8 Discussion

In terms of Mean Absolute Error, the proposed algorithm is as accurate as polynomial fitting within the bandwidth of the reference lamp. Outside of this band third-order fitting has errors that are 12-35 times higher, while our algorithm has only 10-30% greater error.

It is difficult to directly compare the errors reported in previous papers on wavelength calibration accuracy. The main reason for this is that the various spectrometers that were used in other studies have highly varying wavelength resolutions due to different properties in terms of slit width, focal length, grating period, system distortion, and camera pixel size and noise characteristics. For this reason, we have chosen to compare the performance of the proposed algorithm directly with first-, second, and third order polynomial fitting rather than attempt to cross-compare with other studies. As an example, the (all-peaks) standard deviation error for second-order polynomial fitting over ten neon spectra reported in one of the most cited papers [62] is given as 0.005 nm. The spectrometer used in that paper was a Czerny-Turner spectrometer with a reflection grating with a higher resolution than the one used in this chapter (focal length 0.64m and grating 1800 lines/mm). The most similar result for our chapter (second-order fitting, all-peaks, 1000 lines/mm) has $\overline{SD} = 0.00596$ nm, taken over 100 neon spectra. For the two most similar methods in the literature that wavelength calibrate using a physical model, much smaller accuracy is reported: in Ref. 79 an accuracy of 0.1nm is reported. However, for this case the resolution of the 130mm focal length spectrometer is signifcantly less than that of our own systems and is reported to be 0.5 nm at the central wavelength and up to 2 nm at the edge wavelengths. In Ref [80], the Czerny-Turner monochromator had a focal length of 300mm, a grating density of 1200 lines/mm, and a 2160-pixel linear CCD detector with 14 μm pixel size. The authors report a mean absolute value error of 0.16 nm, and a standard deviation of error of 0.22 nm.

One should note that polynomial fitting algorithms are connected to the instrument modeling approaches because *sine* and *cosine* can be approximated as series expansions,

and the terms of those expansions are closely approximated by cubic polynomials. This explains why the accuracy of third-order fitting and the proposed algorithm are similar for all cases within the region of the reference lines as evidenced by the leave-one-out evaluation.

4.9 Conclusion

In this chapter a novel wavelength calibration algorithm is proposed, which outperforms traditional polynomial fitting based methods, particularly in spectral bands that lie outside of the range of spectral lines provided by the reference lamp. Our method was demonstrated to be between 12-121 times more accurate that third-order fitting in such bands when compared to third-order fitting, and 2.5-6 times more accurate than second-order fitting. When compared to other recently proposed wavelength calibration algorithms that make use of a physical model of the system, the proposed algorithm is significantly faster and simultaneously fits to a larger range of physical parameters in the system, including distortion of the image plane. This is achieved by performing linear regression within the brute force search for those parameters which linearly relate wavelength and pixel position on the detector.

A secondary, but nevertheless important, contribution in this chapter is the introduction of a number of new evaluation methods for wavelength calibration accuracy. The traditional approach of evaluating error by inspecting each peak in the reference spectrum (ALL) is augmented with two approaches borrowed from chemometrics: Leave-one-out-cross-validation (LOO) and leave-half-out (LHO) evaluation. The former involves performing wavelength calibration using all but one of the reference peaks, and subsequently calculating error for that one peak. The same process is repeated for each peak. In this way the error of wavelength calibration for peaks within the spectral range of the lamp is better estimated since the peaks that are inspected were not part of the calibration process. LHO on the other hand provides a better estimate of accuracy outside of the range of spectral peaks in the reference lamp by using only one half side of the spectral lines for calibration, and the other half to calculate error. We believe that

these metrics should become the standard in evaluating wavelength calibration going forward.

In terms of future work we believe there is scope to improve the proposed algorithm by obtaining a better first guess of the core spectrometer parameters in the search algorithm. This could be achieved by using the approach of Ref [78] in which a set of simultaneous equations can be derived from the physical model to approximately solve for these parameters. Further we believe better accuracy could be obtained if the spectral line positions for the reference lamp were corrected to account for the refractive index of air [83, 85] as has been done for other wavelength calibration methods; we made no attempt to do this in this chapter.

In the next chapter, the proposed algorithm is adapted for direct wavenumber calibration of Raman spectrometers using a set of Raman wavenumber standards. The new metrics are also used to evaluate the approach, this time being applied to wavenumber values for the peaks in the Raman spectra.

4.10 Appendix

In this appendix evaluation of the proposed algorithm is shown for the error metric of standard deviation (Fig. 4.7) and Root Mean Square Error (Fig. 4.8) as defined in Section 4.6.2. These results correspond to those shown in Fig. 4.6 in the main body of the paper for the case of the error metric Mean Absolute Error. These additional results are shown here to help in comparing with results from other papers.

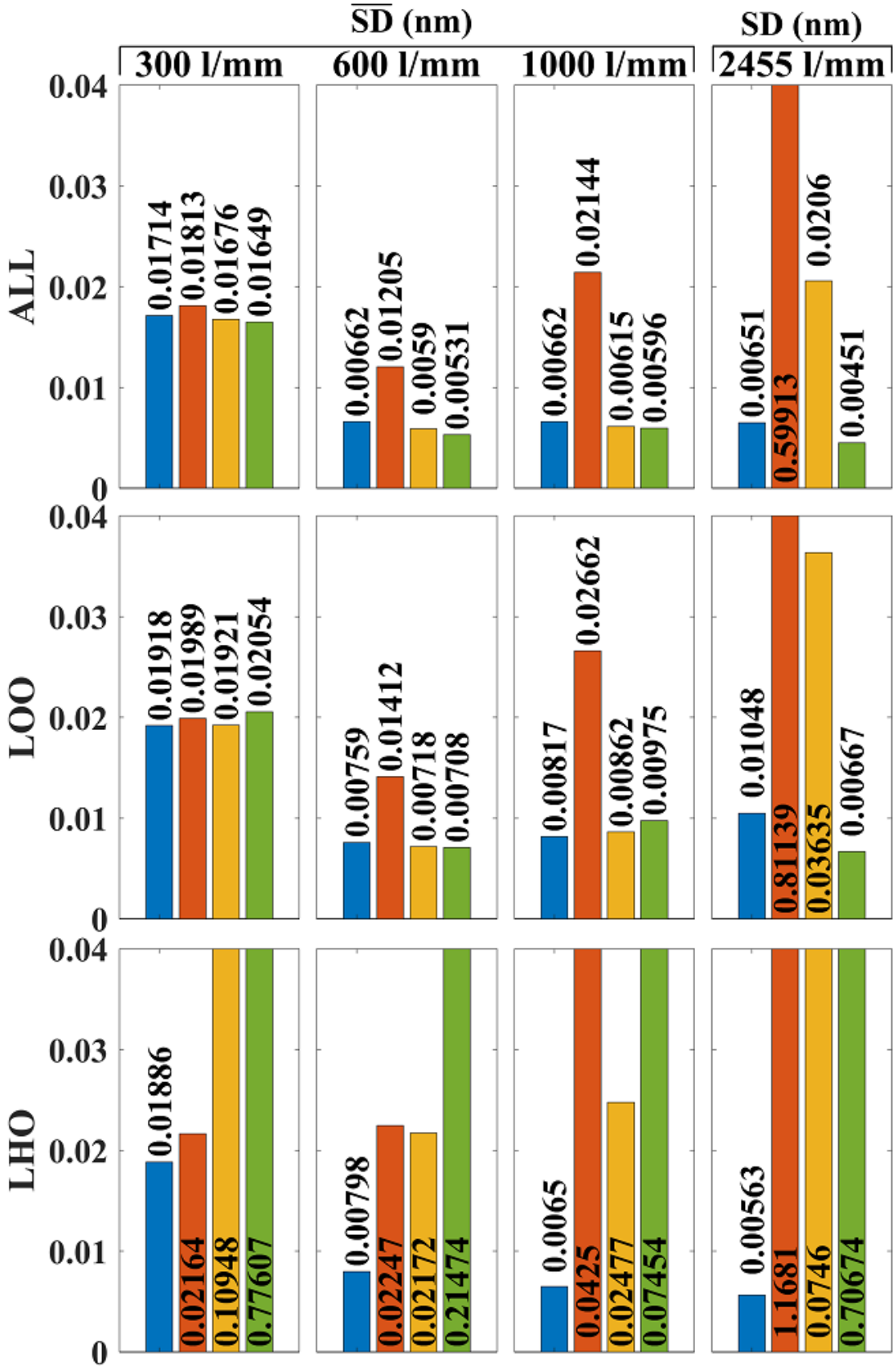


Figure 4.7: Standard Deviation result for system

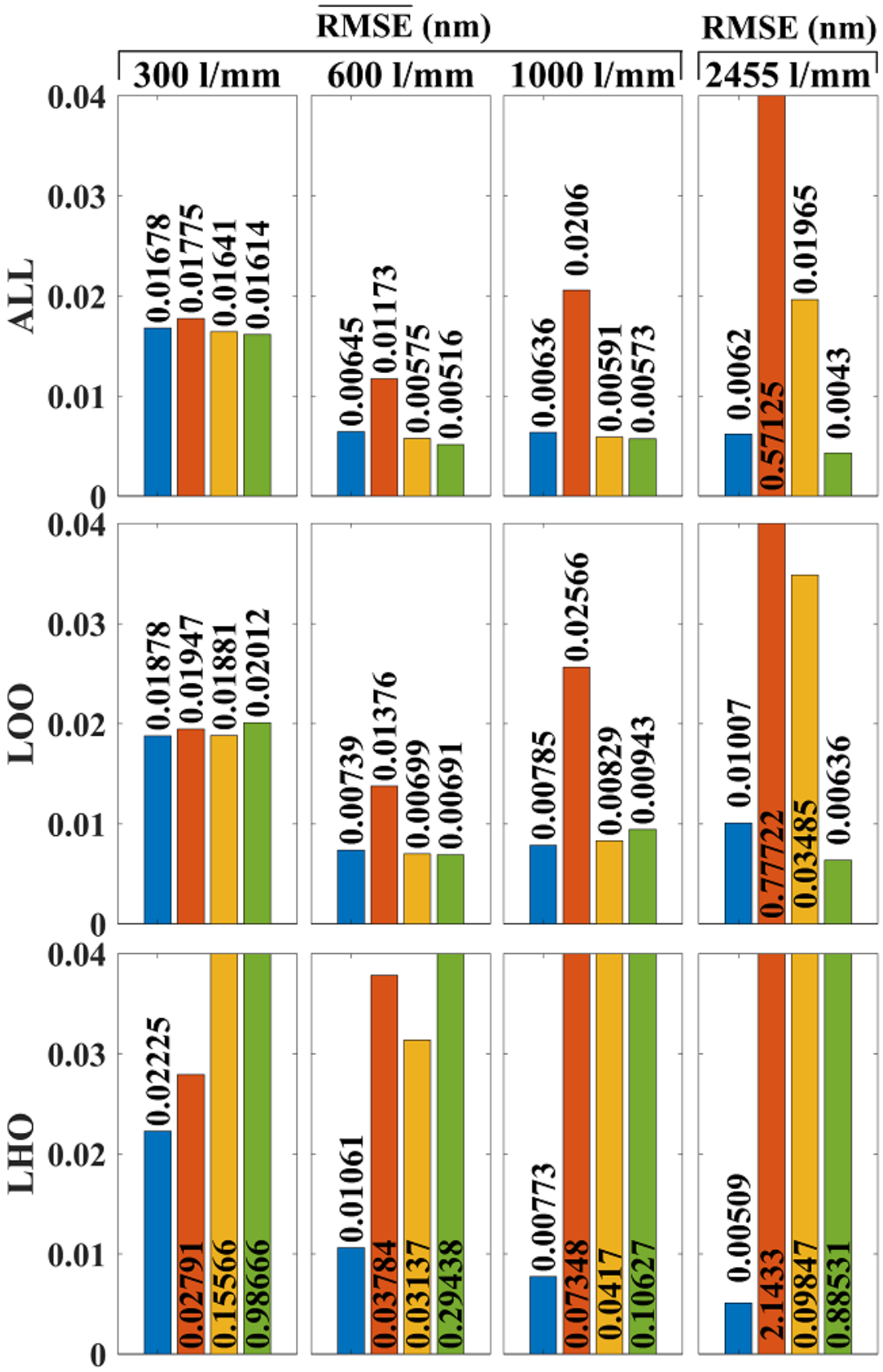


Figure 4.8: RMSE result for system

Chapter 5

Wavenumber Calibration by Modelling the Raman Spectrometer

The work in this chapter has been submitted to the Journal of Raman Spectroscopy with the following reference: ***Liu, Dongyue, and Bryan M. Hennelly. "Wavenumber Calibration by Modelling the Raman Spectrometer." submitted to Journal of Raman Spectroscopy Sept 2022*** **with the following abstract:**

"A direct wavenumber calibration protocol is proposed that replaces polynomial fitting to relate the detector axis and the wavenumber axis. The physical model of the Raman spectrometer is used to derive a mathematical expression relating the detector plane to the wavenumber axis, in terms of the system parameters including the spectrograph focal length, the grating angle, and the laser wavelength; the model is general to both reflection and transmission gratings. A fast search algorithm detects the set of parameters that best explains the position of spectral lines recorded on the detector for a known reference standard. Using three different reference standards, four different systems, and hundreds of spectra recorded with a rotating grating, we demonstrate the superior accuracy of the technique, especially in bands outside of the outermost reference peaks when compared with polynomial fitting. Although similar approaches have been investigated for wavelength calibration, we believe this is the first paper to investigate direct wavenumber calibration based on physical modeling. We also provide a thorough

review of wavenumber calibration for Raman spectroscopy and we introduce several new evaluation metrics to this field borrowed from chemometrics, including leave-one-out and leave-half-out cross validation."

5.1 Introduction

The overall goal of this thesis is to develop a set of calibration tools that can help move Raman spectroscopy from the research lab into the clinic. These calibration tools can be classified as (i) wavenumber calibration tools and (ii) intensity calibration tools. As for the previous chapter, it is wavenumber calibration that is the subject of this chapter. In the previous chapter, we investigated a novel method of wavelength calibration protocol making use of an atomic emission reference spectrum such as neon. Wavelength calibration can be used as the first in two steps to implement wavenumber calibration; however, as described below, the second step of measuring the laser excitation wavelength and using this to perform wavenumber conversion on the calibrated wavelength axis is not always trivial. Arguably it is preferable to apply direct wavenumber calibration by using a Raman wavenumber reference such as 4-acetamidophenol or the polymer that was investigated in Chapter 3. In this chapter, the wavelength calibration protocol developed in the previous chapter is used as the basis for a novel direct wavenumber calibration protocol.

In Section 2.5 a thorough literature review of wavenumber calibration was provided. In that review we provided a detailed review of direct wavenumber calibration in the literature, compare with the alternative approach of wavelength calibration followed by wavenumber conversion. In this review we have also compared the various contributions over several important features including: the reference materials used, the number of peaks in the reference spectrum, the resolution of the systems, methods for sub-pixel interpolation, and calibration accuracy. All of the literature to date on the subject of direct wavenumber calibration employs a low order polynomial (typically of order 3) to fit the detector pixel and reference wavenumber shift pairs that are recorded from a reference standard; this polynomial provides the calibrated wavenumber axis. In

this chapter, we propose an alternative to polynomial fitting, which provides superior accuracy and precision, particularly in bands outside of the outermost peaks in the wavenumber reference spectrum. The method is based on deriving the relationship between wavenumber and detector pixel for an arbitrary Raman spectrometer based on the physical model, which is defined in terms of the system parameters including the spectrometer focal length, grating angle etc. A search algorithm estimates the set of parameters that are optimal in terms of fitting the detector pixel and reference wavenumber shift pairs. The method is tested on hundreds of spectra recorded using four different systems with varying resolution including a reflection Czerny-Turner spectrometer with a motorised grating as well a low f-number spectrometer with a holographic transmission grating. In all cases, it is shown that the method is superior to polynomial fitting and we believe that this method could be considered for inclusion in future iterations of ASTM-E1840. [2]

The breakdown of this chapter is as follows: In Section 5.2, a physical model of the general Raman spectrometer is analysed and the relationship between wavenumber and detector pixel is derived. We note that this naturally builds upon the approach taken in the previous chapter in which the wavelength axis and pixel axis were related using the same physical model; in this chapter we also investigate what polynomial order would best fit this relationship, which has been a subject of debate in previous papers. Based on this relationship, an algorithm is proposed in Section 5.3 to replace polynomial fitting. In Section 5.4, the methods are detailed including a description of the overall calibration routine, and the metrics used for accuracy/precision. In Section 5.6, results are presented, which are followed by a brief conclusion.

5.2 Relationship between wavenumber and pixel-position in a spectrometer

5.2.1 Relationship between wavelength and pixel position for a generalized spectrometer with a rotating grating

Here we simply summarise the result obtained in Section 4.3 in the previous chapter, which derived the relationship between the wavelength of a point-source at the spectrometer slit, to the position of the image of this point on the array detector. We recall that the derivation is general for both transmission and reflection gratings, and the direct wavenumber calibration algorithm that builds on this and which appears in the next section can, therefore, be applied to Raman spectrometers that employ both types of gratings. The relationship between the wavelength, λ, and the position, x, on the detector plane is given by:

$$x = \frac{f}{T}\tan\left\{\theta_d + \sin^{-1}\left[\frac{n\lambda}{d'} - k\sin\left(-\alpha-\theta_d\right)\right] - \alpha\right\} + \frac{C}{T} \tag{5.1}$$

where n is the diffraction order, f is the focal length of the spectrometer, T is the pixel pitch of the detector, θ_d represents the angle of the grating, k is an integer with value -1 for a transmission grating and $+1$ for a reflection grating, α is half the deviation angle of the spectrometer, and C represents misalignment of the centre of the detector array with respect to the optical axis. Curvature of the slit image in the detector plane is often caused by the displacement of the irradiance spot vertically along the slit resulting in an oblique angle of the light, γ, incident on the grating [81, 82, 134, 138], and $d' = d\cos\gamma$, where d is the grating period. We refer the reader to Fig. 4.1 in the previous chapter for an illustration of the rotating grating, which includes each of these parameters.

5.2.2 Relationship between wavenumber and pixel position for a generalized spectrometer with a rotating grating

The wavenumber conversion formula is defined in Equation 5.2 below:

$$\upsilon = \left(\frac{1}{\lambda_L} - \frac{1}{\lambda}\right) 10^7 \tag{5.2}$$

where the wavenumber, υ, is defined in units of cm^{-1} and wavelength is defined in terms of nm and λ_L denotes the laser wavelength. To define the relationship between wavenumber and pixel position this wavenumber conversion formula is applied to Equation 5.1. The forward and inverse relations are defined in Equation 5.3 and Equation 5.4 as follows:

$$x = \frac{f}{T}\tan\left\{\theta_d + \sin^{-1}\left[\frac{n\lambda_L}{d'(1-10^2\nu\lambda_L)} - k\sin(-\alpha-\theta_d)\right] - \alpha\right\} + \frac{C}{T} \tag{5.3}$$

$$\upsilon = \frac{-10^{-2}n}{d'\left\{\sin\left[\tan^{-1}\left(\frac{xT-C}{f}+\alpha-\theta_d\right)\right] + k\sin(-\alpha-\theta_d)\right\}} + \frac{10^{-2}}{\lambda_L} \tag{5.4}$$

In these two equations, the position on the detector, x, is defined in units of pixels. Equation 5.4 is used as the basis of the wavenumber calibration algorithm that is outlined in Section 5.3. Before deriving this algorithm, we first explore the relationship between x and υ for two different Raman spectrometers, which are illustrated in Fig. 5.1, and which are later used for experimental validation of the proposed algorithm. The purpose here is to examine the non-linearity of this relationship for different systems, in an effort to elucidate the inconsistent results presented in the literature to date on the optimal polynomial order to best relate x and υ. Full details of these two spectrometers are provided in the previous chapter in Section 4.3. For the purpose, of this chapter, it suffices to point out some key details and to provide the set of parameters for each system in Table 5.1. For ease of reading, there is some repition below in terms of the description of the spectrometers that appeared in Section 4.3.

The Czerny-Tuner spectrograph (Andor Shamrock 500; SR-500i-A; Andor UK) is

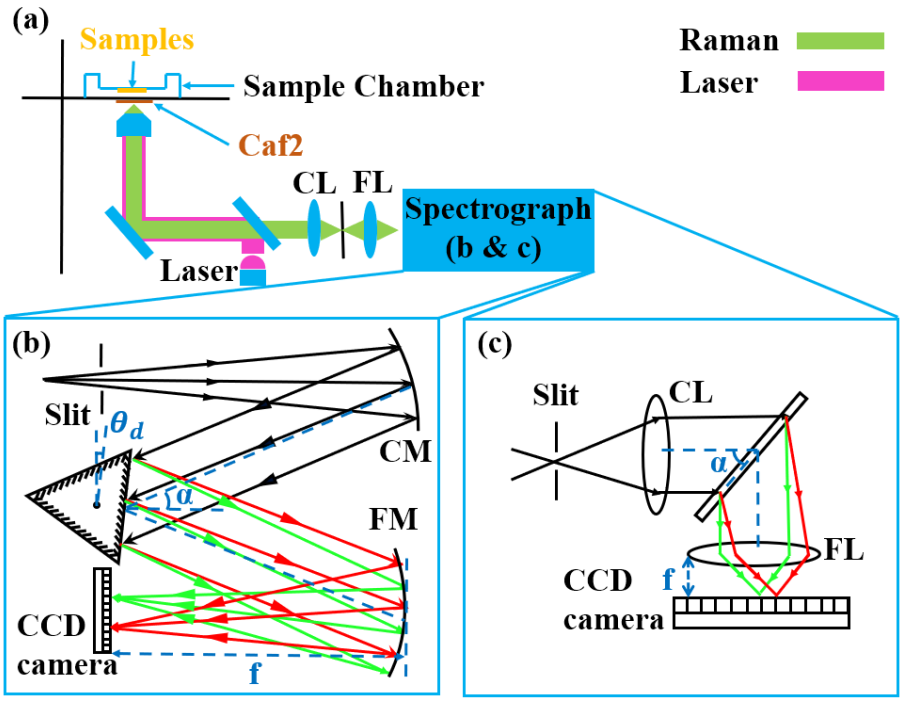


Figure 5.1: (a) Basic Raman spectrometer with epi-illumination; (b) A Czerny-Tuner spectrograph with a rotating grating; the parameters shown in the illustration appear in Equation 5.1 in the text; (c) A transmission spectrograph using a holographic grating. Both types of spectrographs are used in this study.

illustrated in Fig. 5.1(b); it used parabolic mirrors with focal length of 500mm and contains a motorized rotating grating with interchangeable reflection gratings of period, 300, 600, and 1000 lines-per-mm. A transmission spectrometer (HOLOSPEC-F/1.8I-VIS; Andor, UK) is also investigated in this study, illustrated in Figure 5.1(c). This lens based system used volume-phase holographic transmission grating with 2455 lines/mm. A same detector is used for both cases: a cooled CCD (Andor iDus; DU420A-BR-DD; Andor UK) with 256×1024 pixels with a pixel-pitch, T of 26 µm Further details on both spectrographs are available in the previous chapter in Section 4.3. The set of parameters that describe the Raman spectrometers are provided in Table 5.1. These parameters are required for the algorithm that is proposed in Section 5.3.

In order to elucidate the nonlinear relationship between wavenumber and pixel position for these Raman spectrometers, Equation 5.4 is plotted for integer values of x in the range $-N/2 \rightarrow N/2-1$, as shown in Fig. 5.2. The wavenumber values have been normalised for comparison in the figure. It is clear that the Raman spectrometer containing the transmission spectrograph with the short focal length and high dispersion

Parameter	Unit	Czerny Turner	Transmission
Reflection/ transmission (k)	NA	+1	-1
Diffraction order (n)	NA	-1	+1
Grating Period (d)	lines/mm	300 600 1000	2455
Half the deviation angle (α)	degrees	10.94	45
Grating angle (θ_d)	Degree	4.8 10.2 17.2	0
Focal length (f)	mm	500	85
Bandwidth	cm^{-1}	~4190.78 ~2696.53 ~1650.80	~2430
Average resolution	cm^{-1}	9.73 5.05 3.19	5.51
Camera pixel pitch (T)	µm	26	26
Camera width (N)	pixels	1024	1024
Camera centre position (C)	pixels	0	0
Laser wavelength (λ_L)	nm	532	532

Table 5.1: The parameters for the two spectrometers illustrated in Fig. 5.1, which are investigated in this study.

grating (2455 lines/mm) exhibits the most non-linear relationship over the span of the detector. Interestingly, for the spectrometer with the Czerny-Turner spectrograph, the most linear profile belongs to the case of the 1000 lines/mm grating, and the profile becomes more linear as the grating period reduces.

These four profiles (without wavenumber normalisation) were subject to polynomial fitting with orders from 1-7 and the mean absolute error in wavenumber between the resultant polynomials and the profile were calculated. These values are presented in Table 5.2. As polynomial order increases, the fitting error reduces; however, this trend is different for all four cases and depends on the non-linearity of the $x-v$ relationship. The polynomial order to use in direct wavenumber calibration has been the subject of debate in the literature. [3, 39, 44, 49, 69, 91, 93, 99, 100, 140] with various orders being suggested as optimal for different cases. The polynomial order must be low in order to avoid high error in bands outside of the outermost spectral lines in the reference spectrum, and there is general consensus in recent literature that an order 3

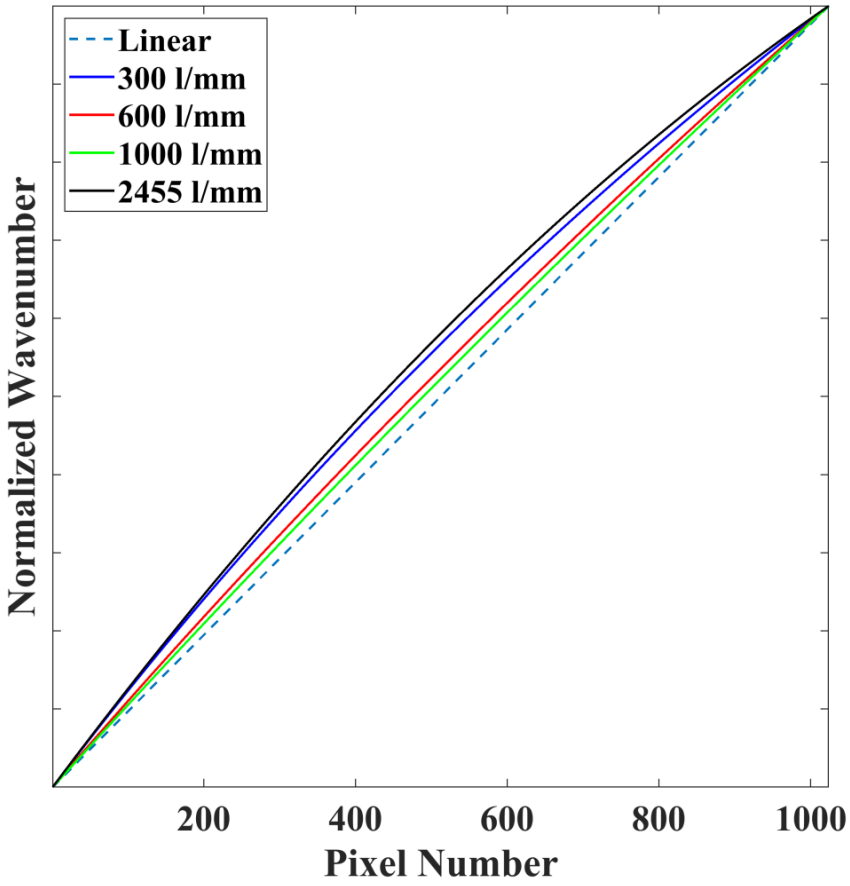


Figure 5.2: Investigation of the non-linearity of the (x, ν) relationship for the different spectrometers. Wavenumber values are normalised over the span of the detector for comparison.

is preferred. [39, 44, 97]. For this order, we predict a theoretical limit of mean absolute errors from 0.0043-0.49 depending on which Raman spectrometer is calibrated. It can be concluded that the accuracy of direct wavenumber calibration using low order polynomial fitting will be highly variable depending on the $x-\nu$ relationship for a given system, which is defined by the system parameters such as the grating period and focal-length. It should also be noted that this analysis is performed with the assumption of negligible optical distortion, which may place a further limit on accuracy. Although the $x-\nu$ relationship cannot the theoretically perfectly modeled by a low order polynomial, this is not so for the $x-\lambda$ relationship, which can be ideally modelled by a second order polynomial [78, 89] in the absence of distortion as described in Section 4.3, see Fig. 4.3. This may strengthen the argument that wavelength calibration followed by wavenumber conversion provides superior accuracy [62] in general; however, recent experimental

results suggest approximate parity between both approaches as reviewed in Section 2.5.

Polynomial Order	300 l/mm	600 l/mm	1000 l/mm	2455 l/mm
1	93.1	27.6	10.4	40.9
2	6.8	1.1	2.2E-1	9.8E-1
3	4.9E-1	3.9E-2	4.3E-3	3.2E-2
4	3.6E-2	1.5E-3	8.7E-5	2.0E-3
5	2.6E-3	5.5E-5	1.8E-6	5.6E-5
6	1.9E-4	2.1E-6	3.6E-8	5.6E-6
7	1.4E-5	7.7E-8	7.2E-10	1.1E-7

Table 5.2: The Mean Absolute Error in units of cm^{-1} following the fitting of the profiles shown in Fig. 5.2 with polynomials of orders 1-7. In this calculation, the wavenumber values have not been normalised.

5.3 Algorithm

In this section, an algorithm is developed that relates the detector pixels, x, to the corresponding spectral wavenumber values, υ, which replaces the step of polynomial fitting in traditional direct wavenumber calibration algorithms. This algorithm is based on Equation 5.4, which mathematically relates x and υ using the physical model of the system. Like traditional polynomial fitting, this algorithm requires a set of matching sub-pixel positions, X_0 and reference wavenumber values V_0, on which to apply the fitting algorithm. Following the recording and processing of a reference wavenumber spectrum, such a set of matching pairs will be available. Explicitly, these are defined as follows:

$$\begin{aligned} X_0 &= [x_1, x_2, \cdots, x_M] \\ V_0 &= [\upsilon_1, \upsilon_2, \cdots, \upsilon_M] \end{aligned} \tag{5.5}$$

where $x_{i=1:M}$ are the sub-pixel positions of the spectral lines, which have known reference wavenumber values $\upsilon_{i=1:M}$, where there are M useful spectral lines in the reference spectrum.

Equation 5.4 contains several system parameters that define the system and knowledge of their precise values enables accurate fitting of Equation 5.4 to the available data points X_0, V_0. The algorithm searches over a range of values of these system parameters to provide optimal fitting. The set of relevant system parameters, S is given by Equation 5.6.

$$S = [f, T, C, d', \alpha, \theta, \lambda_L] \tag{5.6}$$

Each of these parameters is approximately known based on manufacturer specification or approximate measurement. However, the precise values cannot easily be determined at the outset. Even the effective detector pixel size could be slightly reduced by a small out-of-plane tilt of the detector. To facilitate the discussion that follows, Equations 5.3

and 5.4 are rewritten using operator notation as follows:

$$\begin{aligned} x &= model(S, v) \\ v &= model^{-1}(S, x) \end{aligned} \tag{5.7}$$

The goal is to find values of S that 'best' match X_0 and V_0. One naive approach is to perform a brute-force search over all seven parameters using their approximate values as a starting point; however, this approach is computationally intractable. It is clear in Equation 5.3 that parameters C, T, and f are linear in effect, and can, therefore, be accounted for using linear regression techniques. It is necessary to brute-force search only over a reduced set of four parameters $S' = [d', \alpha, \theta, _L]$, which significantly reduces the scope of the search. The algorithm is defined as follows:

1. The first step provides initial estimates of the key parameters in S, which are defined as $S_0 = [f_0, T_0, C_0, d'_0, \alpha_0, \theta_0, \lambda_{L0}]$ as follows: The values of $(f_0, T_0, d'_0, \lambda_{L0})$ can be taken from the manufacturers specifications for the spectrograph detector, and laser. The value for α_0 is approximately measured manually and both C_0 and θ_0 are estimated using a brute-force search over only these two variables for a single peak pair (x_i, v_i).
2. A brute-force search is performed over $[\alpha, d', \theta, \lambda_L]$ in a small range, centred at $[\alpha_0, d_0, \theta_0, \lambda_{L0}]$ and using values f_0, T_0, C_0 in order to provide an intermediate result:

$$X_j = model([C_0, f_0, T_0, \alpha_j, d'_j, \theta_j, \lambda_{Lj}]) \tag{5.8}$$

   For each unique set of values $[\alpha_j, d'_j, \theta_j \lambda_{Lj}]$ over the search range, the resultant values X_j are linear-regressed with respect to positions at which the peaks were detected, X_0, in order to account for errors in f, T, and C, which provides an updated set of estimated positions, X'_j. For simplicity, we describe this operation in terms of the Matlab functions, `polyfit` and `polyval`, which are used to

implement it as follows:

$$\begin{aligned} P_j &= \texttt{polyfit}(X_j, X_0, p) \\ X'_j &= \texttt{polyval}(X_j, P_j) \end{aligned} \tag{5.9}$$

where the function `polyfit` returns the coefficients of polynomial degree $p = 1$ that is the best fit (in a least-square sense) to describe the transformation between X_j and X_0. This is followed by the function `polyval`, which applies this transformation to X_j using these coefficients in order to provide the updated values for X'_j.

The specific set of parameters, $S_{min} = [f_0, T_0, C_0, \alpha_{min}, d'_{min}, \theta_{min}, \lambda_{min}]$, and linear regression coefficients defined by P_{min}, which produce the set X_{min} that most closely match the actual pixel values X_0 are taken to be the system parameters. This is determined by minimising the error function defined in Equation 5.10:

$$err = \sum_{i=1}^{N} (x_i^j - x_i)^2 \tag{5.10}$$

where $X'_j = [x_1^j, x_2^j, \cdots, x_M^j]$. We acknowledge that a tight-grid brute force search over four parameters would not normally be applied since modern search algorithms are far more efficient such as steepest descents and simplex searching. However, for the purpose of this chapter, a brute-force search was sufficient.

3. Now that the system parameters S_{min} and P_{min} are known, the integer pixel (centre) positions can be related to the corresponding wavenumber values, V, thereby providing wavenumber calibration for the spectrograph. In this final step, the pixel positions are projected into the wavenumber domain by using the opposite process outline in Step 2. Taking the centre of the CCD pixels to be given by $X_{CCD} = [1, 2, \cdots, 1024]$, the matching wavenumber values are given by:

$$V_{CCD} = model^{-1}[S_{min}, \texttt{polyval}(X_{CCD}, P_{min})] \tag{5.11}$$

The algorithm is general for any spectrometer; however the constant integer values

of n and k that are used in Equation 5.4 should be chosen accordingly. In our experiments, the transmission spectrometer uses the $n = 1$ diffraction order, while for the Czerny-Turner system uses the $n = -1$ diffraction order. The value of k depends on the use of a reflection or transmission grating. Therefore, for the case of the transmission spectrometer $k = -1$, while for the Czerny-Turner system $k = 1$ as shown in Table 5.1.

5.4 Overall calibration procedure

The overall calibration protocol is similar to that outlined in Algorithm 2 in the previous chapter, see Section 4.4.2 with some additional steps. Step 1 is to record a Raman spectrum from a reference material containing some number of sharp, symmetrical and well defined, known peak wavenumbers. Here, we use three reference materials: 4-acetamidophenol (Sigma, Ireland), benzonitrile (Sigma, Ireland), and a commercial polymer (Ibidi, GmbH). Recorded spectra are shown in Fig. 5.3, which also illustrates the different bands that were recorded using the four different spectrometers.

The wavenumber values of the lines that are highlighted in Fig. 5.3 are provided in Table 5.3. These reference values and uncertainties have been taken from ASTM [2] and Ref. 3. For the case of the 300 lines/mm grating all of the values shown in Table5.3 were used for wavenumber calibration and a reduced set were used for the other gratings as detailed in the caption for Table 5.3. Step 2 is to identify the sub-pixel position of the lines in the recorded spectrum. Various methods are reviewed in Section 2.5 on how this can be achieved; here, we fit a Lorentzian function to the intensities of the pixels in the region of the peak of the following form: [139]

$$\frac{P_1}{(x-P_2)^2+P_3}+P_4 \tag{5.12}$$

where P_1, P_2, P_3 and P_4 are the fit-parameters. The value of P_2 is taken to be the sub-pixel position of the peak. An example of this approach is given in Fig. 5.3 (d) in which we show a Lorentzian function that has been fit to one peak in the benzonitrile spectrum. In the results section below, all of the peak positions in each reference spectrum are estimated with sub-pixel accuracy using this approach. Step 3 is the application of the algorithm described in Section 5.3 using these sub-pixel positions and the matching wavenumber values given in Table 5.3.

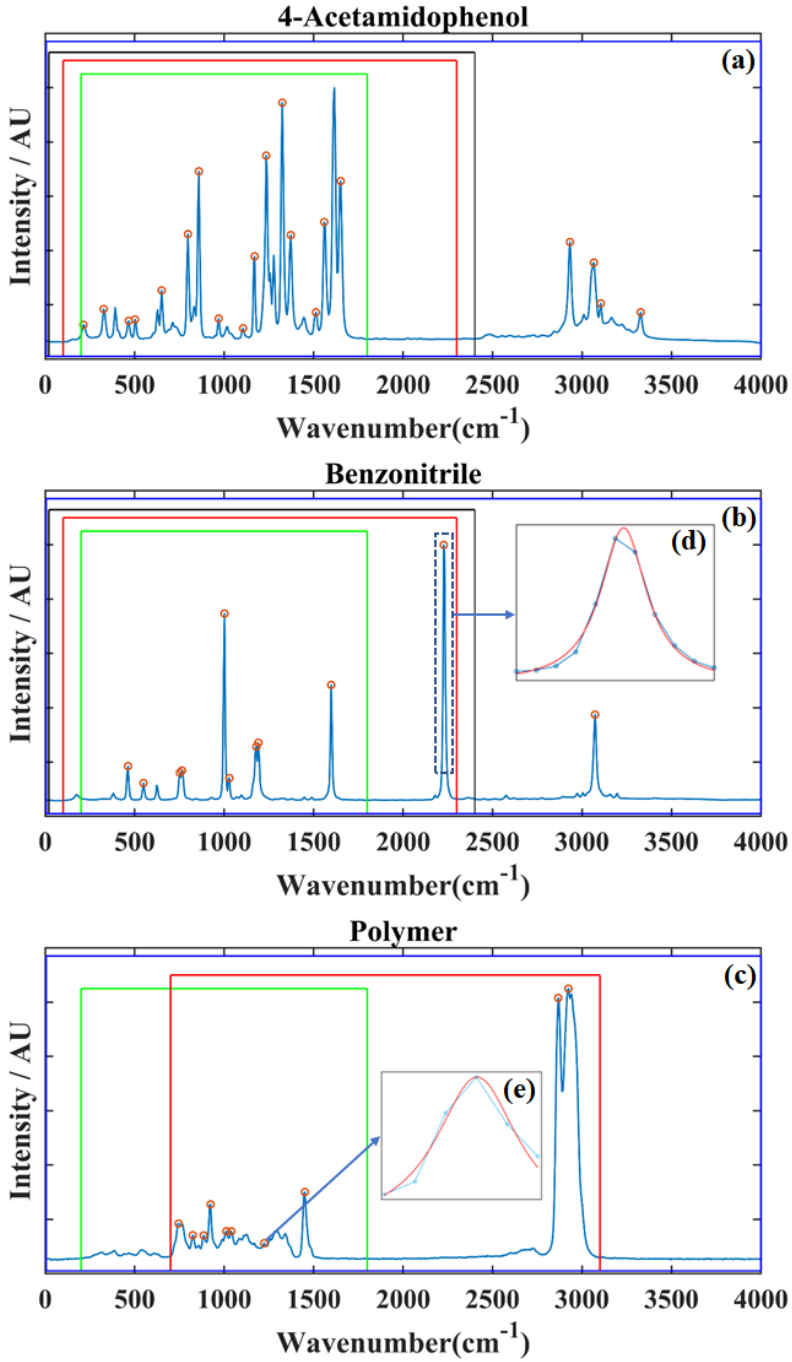


Figure 5.3: Sample spectra recorded from the three different reference materials. The spectrum of (a) 4-acetamidophenol, (b) benzonitrile and (c) commercial grade polymer. The recorded bands using the four spectrometers are highlighted in different colour boxes: The black area corresponds to the 2455 lines/mm grating; green is 1000 lines/mm; red 600 lines/mm; blue 300 lines/mm. (d) A single peak from the benzonitrile is expanded. A Lorentzian function is fitted to the data points around the peak in order to detect the peak centre with sub-pixel accuracy as described in Section 5.4. (e) A further example is given of Lorenzian peak fitting, this time applied to a relatively broader and weaker peak.

Wavenumber (cm^{-1}) ± Standard Deviation		
4-acetamidophenol		
213.3±1.77	329.2±0.52	465.1±0.30
504±0.60	651.6±0.50	797.2±0.48
857.9±0.50	968.7±0.60	1105.5±0.27
1168.5±0.65	1236.8±0.46	1323.9±0.46
1371.5±0.11	1515.1±0.70	1561.5±0.52
1648.4±0.50	2931.1±0.63	3064.6±0.31
3102.4±0.95	3326.6±2.18	
Benzonitrile		
460.9±0.73	548.5±0.82	751.3±0.74
767.1±0.59	1000.7±0.98	1026.6±0.81
1177.9±0.82	1192.6±0.56	1598.9±0.70
2229.4±0.39	3072.3±0.41	
Polymer		
743.5±0.56	828.0±0.90	886.6±0.54
923.1±0.15	1005.7±0.17	1041.7±0.68
1224.6±0.60	1449.3±0.30	2869.0
2914.0		

Table 5.3: Reference spectral lines and uncertainties used in this chapter. Values for 4-acetamidophenol and benzonitrile are taken from ASTM, [2] and values for the polymer are taken from Reference 3. Different numbers of reference lines were used for the different spectrometers depending on their bandwidth as illustrated in Fig. 5.3. For 4-acetamidophenol: 300 lines/mm-20 peaks, 600 lines/mm-17 peaks, 1000 lines/mm-14 peaks, 2455 lines/mm-14 peaks; benzonitrile: 300 lines/mm-11 peaks, 600 lines/mm-10 peaks, 1000 lines/mm-9 peaks, 2455 lines/mm-10 peaks, polymer: 300 lines/mm-8 peaks, 600 lines/mm-10 peaks, 1000 lines/mm-8 peaks. The uncertainty for the lines of the polymer is based on the four seconds recording in Reference 3. The uncertainty for the latter two lines is not available.

5.5 Experiment design

5.5.1 Recording of reference spectra

The performance of the proposed wavenumber calibration algorithm is examined across two Raman spectrometer designs and four different gratings periods as described in previous section. For the case of the Czerny-Turner system, all three reference materials were investigated, while for the transmission spectrometer the polymer was omitted. For each of the three gratings in the Czerny-Turner spectrometer, and for each of the three materials, 100 different reference spectra are recorded with changes in the grating rotation angle. For each of these cases, a rigorous evaluation of the performance of the calibration is possible by calculating the ensemble average of the error metrics defined below, across the set of 100 reference spectra.

While the polymer slide has the advantage of requiring no preparation whatsoever, the 4-acetamidophenol and benzonitrile are in powder and liquid form, respectively. These were both mounted in an Ibidi chamber slide . The base of the chamber was drilled to create an open aperture, which was sealed using a Raman grade Calcium Fluoride coverslip (Crystran, UK), which produces a negligible Raman spectrum except for a single peak at 321 cm^{-1}.

To minimise the effect of shot noise, the accumulation time was maximised to provide a photon count just less than the saturation level of the CCD. Rather than use Full Vertical Binning, which can produce error in the presence of image distortion as disucssed in the previous chapter, area scan images were recorded by the detector; the row of pixels containing the spectrum was cropped.

5.5.2 Error metrics

Several different error metrics have been reported in the literature (see Table 2.1) including, mean absolute error (MAE), the standard deviation (SD), and the root mean square error (RMSE), all of which are measured in this chapter. These metrics are defined below using the same notation as used in Section 5.3. Initially, we define the

error for a single line at known wavenumber v_i to be given by:

$$error(v_i) = calibrated(v_i) - v_i \tag{5.13}$$

where the v_i value is taken from Table 5.3 and the calibrated value, $calibrated(v_i)$, is taken from the set V_{min}:

$$V_{min} = model^{-1}[S_{min}, \texttt{polyval}(X_{min}, P_{min})] \tag{5.14}$$

In some papers, the absolute error (AE) is reported for one peak or for a range of peaks. More traditionally, the mean error is reported for all of the peaks in the reference spectrum as follows:

$$MAE = \frac{1}{M}\sum_{i=1}^{M}|error(v_i)| \tag{5.15}$$

$$RMSE = \sqrt{\frac{1}{M}\sum_{i=1}^{M}|error(v_i)|^2} \tag{5.16}$$

$$SD = \sqrt{\frac{1}{M-1}\sum_{i=1}^{M}|error(v_i) - \overline{error(v_i)}|^2} \tag{5.17}$$

where $\overline{error(v_i)}$ denotes the mean error. In order to provide a more reliable estimate of the error, the above metrics can be calculated for a set of K different spectra where the grating is moved between captures. The ensemble average of each of the above metrics is calculated over these K reference spectra as follows:

$$\overline{MAE} = \frac{1}{K}\sum_{k=1}^{K} MAE(k) \tag{5.18}$$

$$\overline{RMSE} = \frac{1}{K}\sum_{k=1}^{K} RMSE(k) \tag{5.19}$$

$$\overline{SD} = \frac{1}{K}\sum_{k=1}^{K} SD(k) \tag{5.20}$$

5.5.3 Evaluation methods

Here, we describe a number of evaluation methods, making use of the above metrics, which were proposed for the evaluation of wavelength calibration in the previous chapter; see Section 4.6.2. The latter two are borrowed from the area of multivariate statistical analysis [136, 137] and are used here for the first time (to the best of our knowledge) in the evaluation of wavenumber calibration:

1. *All-peaks*: Here all of the calibrated peaks from the reference are used in the error analysis. This is the typical value reported in the literature to date. Taking a spectrum of the reference sample 4-acetamidophenol as an example, which contains 20 reference lines υ_i for $i = 1 \rightarrow 20$, all 20 values of υ_i and the matching sub-pixel positions are used to perform the given calibration routine. Using the resultant calibrated wavenumber axis, the error function for each υ_i is calculated according to Equation 5.10, which enables the MAE to be calculated according to Equation 5.15, and this procedure is repeated for each of the $K = 100$ spectra in the dataset; finally the mean MAE is calculated using 100 results as defined in Equation 5.18. In total a given calibration routine is applied 100 times to calculate the MAE for ALL peaks.
2. *Leave-one-out-cross-validation*: in order to remove any bias from the reference spectrum, we propose the use of 'leave-one-out' (LOO) cross-validation, whereby one peak is removed from the reference spectrum used in the calibration process. The error metric is then applied only to this peak alone following calibration. This process is repeated for each peak in the spectrum and the average value for all cases is calculated. This method must provide a more accurate estimate of wavenumber accuracy *inside* the bounds of the spectral lines provided by the reference spectrum. Again taking a spectrum of the reference sample 4-acetamidophenol as an example, which contains 20 reference lines, the first line υ_1 is removed from the spectrum and the given calibration routine is applied to the remaining 19 lines υ_i for $i = 2 \rightarrow 20$. The resulting calibrated wavenumber axis is applied to the spectrum and the error function defined in Equation 5.10

is applied only to v_1 to obtain $error(v_1)$. Then v_2 is removed from the dataset used for calibration followed by calculation of only $error(v_2)$. This process is repeated in total 20 times to calculate $error(v_i)$ for all $i = 1 \rightarrow 20$, and then the MAE function can be calculated as defined by Equation 5.15. Therefore, LOO analysis applied to a single 4- acetamidophenol spectrum requires application of the given calibration routine 20 times using a different set of 19 reference lines in each instance. This LOO analysis is then applied to all $K = 100$ spectra in the dataset and the mean MAE is obtained as defined in Equation 5.18. In total, LOO analysis of a particular calibration routine requires 2000 applications.

3. *Leave-half-out (LHO)*: We propose an evaluation based on calibrating using the left-most half of the reference peaks and applying the error metric to the right-most peaks of the calibrated spectrum. This is repeated using the right-most peaks for calibration and the left-most for error calculation. The average of the two values is taken. This provides a more accurate evaluation of the accuracy of the wavenumber calibration *outside* the bounds of the reference spectrum lines. Again taking a spectrum of the reference sample 4-acetamidophenol as an example, which contains 20 reference lines, the 10 left-most lines in the spectrum are removed and only the right most 10 lines, v_i for $i = 11 \rightarrow 20$, are used in the given calibration routine. The resulting calibrated axis is applied to the full spectrum and the error function $error(v_i)$ is calculated only for the left 10 lines $i = 1 \rightarrow 10$. This overall process is then repeated this time using the left 10 lines for calibration and the right ten lines to calculate the error function. Thus, two applications of the given calibration routine, using 10 lines in each instance, will provide the 20 values of the error function, $error(v_i)$ for v_i for $i = 1 \rightarrow 20$, which can then be used to calculate the MAE function in Equation 5.15. This LHO analysis is then applied to all $K = 100$ spectra in the dataset and the mean MAE is obtained as defined in Equation 5.18. In total, LHO analysis of a particular calibration routine requires 200 applications.

5.5.4 Comparison with traditional methods of wavenumber calibration

In all cases, the proposed algorithm is compared with equivalent results from first-order up to a seventh-order polynomial. This analysis relates to the discussion at the end of Section 5.2.2 on the non-linear relationship between x and υ.

5.6 Results

In this section, the results are presented for wavenumber calibration using the proposed algorithm and compared with the corresponding set of results from the first-, through to seventh-order (where possible) polynomial fitting. As outlined in the previous section, these results are broken down into three sets of evaluations, corresponding to 'All-Peaks' (ALL), 'Leave-one-out' cross validation (LOO), and 'Leave-half-out' cross validation (LHO). Furthermore, to facilitate comparison with other papers, which use various metrics, these evaluations are performed using three different metrics: $\overline{MAE}$, $\overline{RMSE}$, $\overline{SD}$ as defined in Equations 5.18, 5.19, and 5.20. For the case of the transmission spectrometer, the grating angle could not be adjusted and so only a single spectrum was available. In this case the error metrics used in the evaluation are: MAE, $RMSE$, SD as defined in Equations 5.15, 5.16, and 5.17. The results for the Mean Absolute Error metrics using the 4-acetamidophenol material are shown below in Fig. 5.4. For the case of ALL evaluation, first order fitting is the worst performer by a wide margin for all four systems, which is due to the highly non-linear $x-\upsilon$ relationship; the most inaccurate system is the 300 lines/mm grating, followed by the transmission 2455 lines/mm grating, which is predicted in Table 5.2. We note that the measured error values differ from those in Table 5.2 due to the lower number of wavenumber values used to calculate the error when compared with the theoretical analysis in Section 5.2.2. For each of the four gratings, all of the other polynomial orders, as well as the proposed algorithm, show similar performance: the accuracy for the 300 lines/mm grating are 0.749-1.772 cm^{-1}; for the 600 lines/mm grating the accuracy is 0.315-0.366 cm^{-1}; for the 1000 lines/mm grating

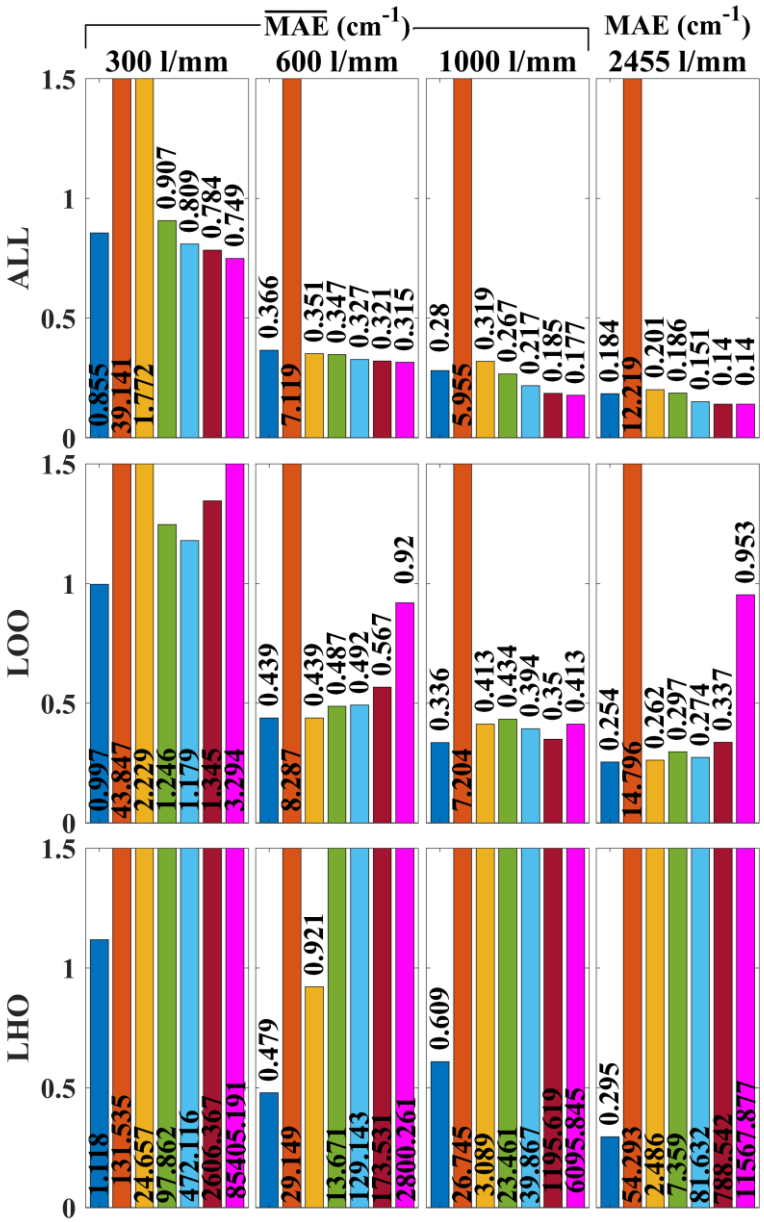


Figure 5.4: Evaluation of direct wavenumber calibration accuracy using Mean Absolute Error applied to 4-acetamidophenol spectra. For the case of the Crezny-Turner reflection spectrometer three different gratings are investigated: 300, 600 1000 lines/mm and for these three cases the $\overline{MAE}$ error metric is applied over 100 spectra with grating movement between capture. The transmission spectrometer with grating 2455 lines/mm is evaluated using a single MAE error metric applied to single spectrum. The results of the algorithm proposed in this chapter are given in blue and the results for first-, second-, third, fourth, fifth, sixth, and seventh-order polynomial fitting are given in orange, yellow, green, blue, red, and pink, respectively. The results of 'All-Peaks' (ALL), 'Leave-one-out-cross-validation' (LOO), and 'Leave-half-out' (LHO) are shown on different rows. For ease of comparison, the same axis range is used for all three evaluations. In several cases, the bars have been capped at 1.5 cm^{-1} to improve visualisation. The correct values are overlaid on the bars in all cases.

the accuracy is 0.177-0.319 cm^{-1}; and for the 2455 lines/mm grating the accuracy is 0.140-0.201 cm^{-1}. Interestingly, the proposed algorithm provides an accuracy equivalent to that of second and third order fitting and the accuracy is improved slightly as the polynomial order is increased. The accuracy does not improve by an order of magnitude, with each increase in polynomial order as predicted by Table 5.2; the limiting factor here is the positional accuracy afforded by sub-pixel interpolation, as well as the resolution of the given system.

It is possible that the slightly better accuracy for order > 3 results from over-fitting of the available data-points. This is suggested by the second evaluation, LOO, which eliminates the possibility of over-fitting; In this case, it is clear that increasing the polynomial order will in general result in increased error *within* the wavenumber band that is bounded by the reference lines. In all cases, the proposed algorithm provided the highest accuracy for LOO evaluation, albeit the error is only slightly lower than for the best polynomial fitting case, which is either the second-, or third-order for each case.

LHO evaluation reveals the strength of the proposed algorithm over traditional methods. This evaluation indicates that in all cases, the proposed algorithm is by far the most accurate in wavenumber bands that are *outside* of the spectral lines in the reference lamp; indeed the accuracy in these bands is only slightly less (0.04-0.273 cm^{-1}) than the accuracy *inside* the bounds according to LOO evaluation. For the case of the 300 lines/mm reflection grating, the proposed algorithm provides the best LHO accuracy, with an error 1.118 cm^{-1} and second-order fitting is next best with an error of 24.657 cm^{-1}; second- and third-order fitting error are 22.1-times and 87.5-times worse than the proposed algorithm, respectively. For the 600 lines/mm reflection grating, the proposed algorithm once again provides the best LHO accuracy with an error of 0.479 cm^{-1}; second-, and third- order fitting provide errors are 1.9-times, and 28.5-times greater. For the third reflection grating of period 1000 lines/mm, the proposed algorithm once again returns the best LHO accuracy with an error of 0.609 cm^{-1}; second-, and third-order fitting provide errors that are 5.1-times and 38.5-times greater. For the transmission grating of period 2455 lines/mm, the proposed algorithm once again returns by far the

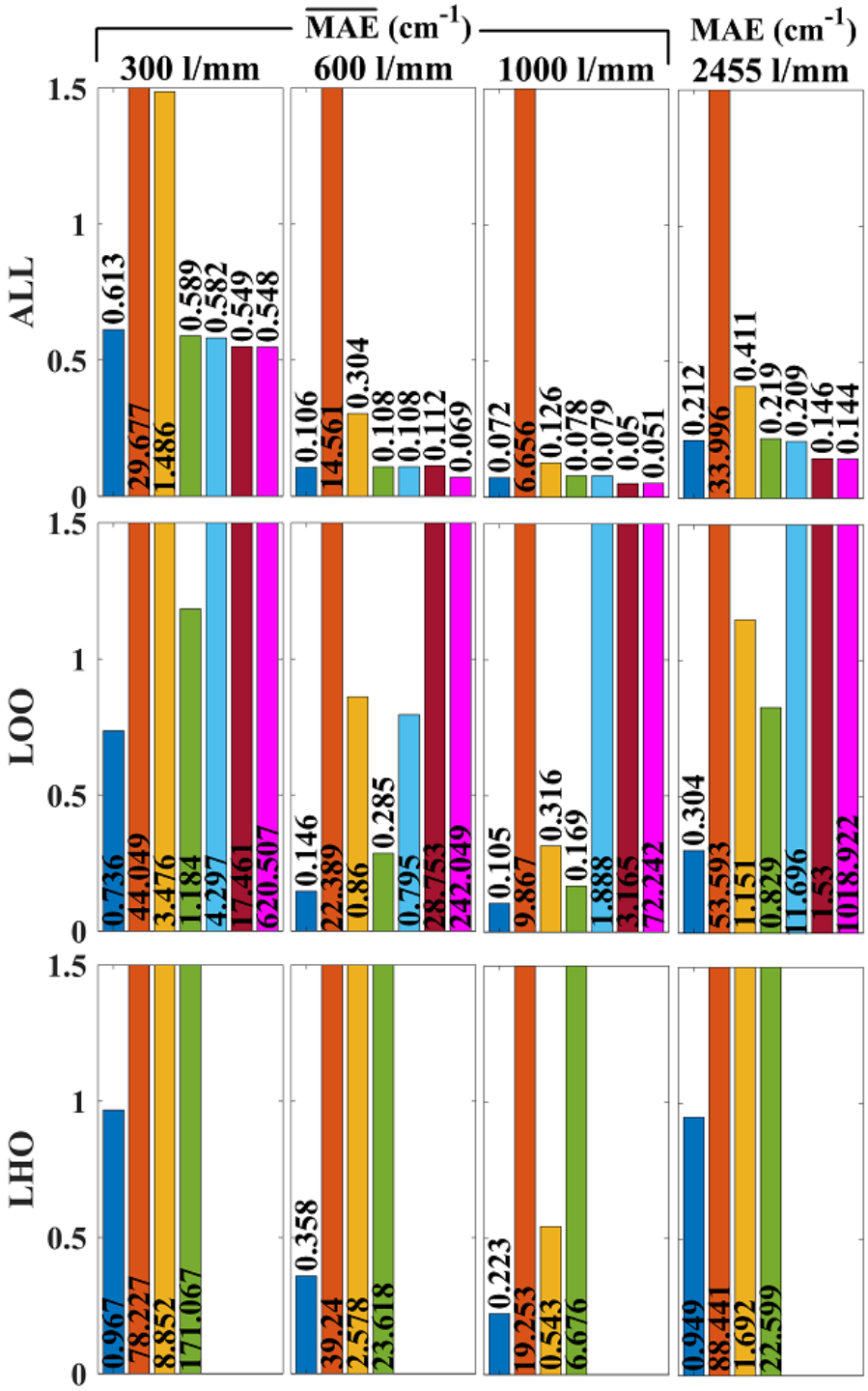


Figure 5.5: Evaluation of direct wavenumber calibration accuracy using Mean Absolute Error applied to benzonitrile spectra. See caption for Fig. 5.4 for further details. Polynomial order >3 could not be applied for LHO evaluation due to lower peak number. The results of the algorithm proposed in this chapter is given in blue and the results for first-, second-, third, fourth, fifth, and sixth-order polynomial fitting are given in orange, yellow, green, blue, red, and pink, respectively.

best LHO accuracy with an error of 0.295 cm^{-1}; second-, and third-order fitting provide errors that are 8.4-times and 24.9-times greater.

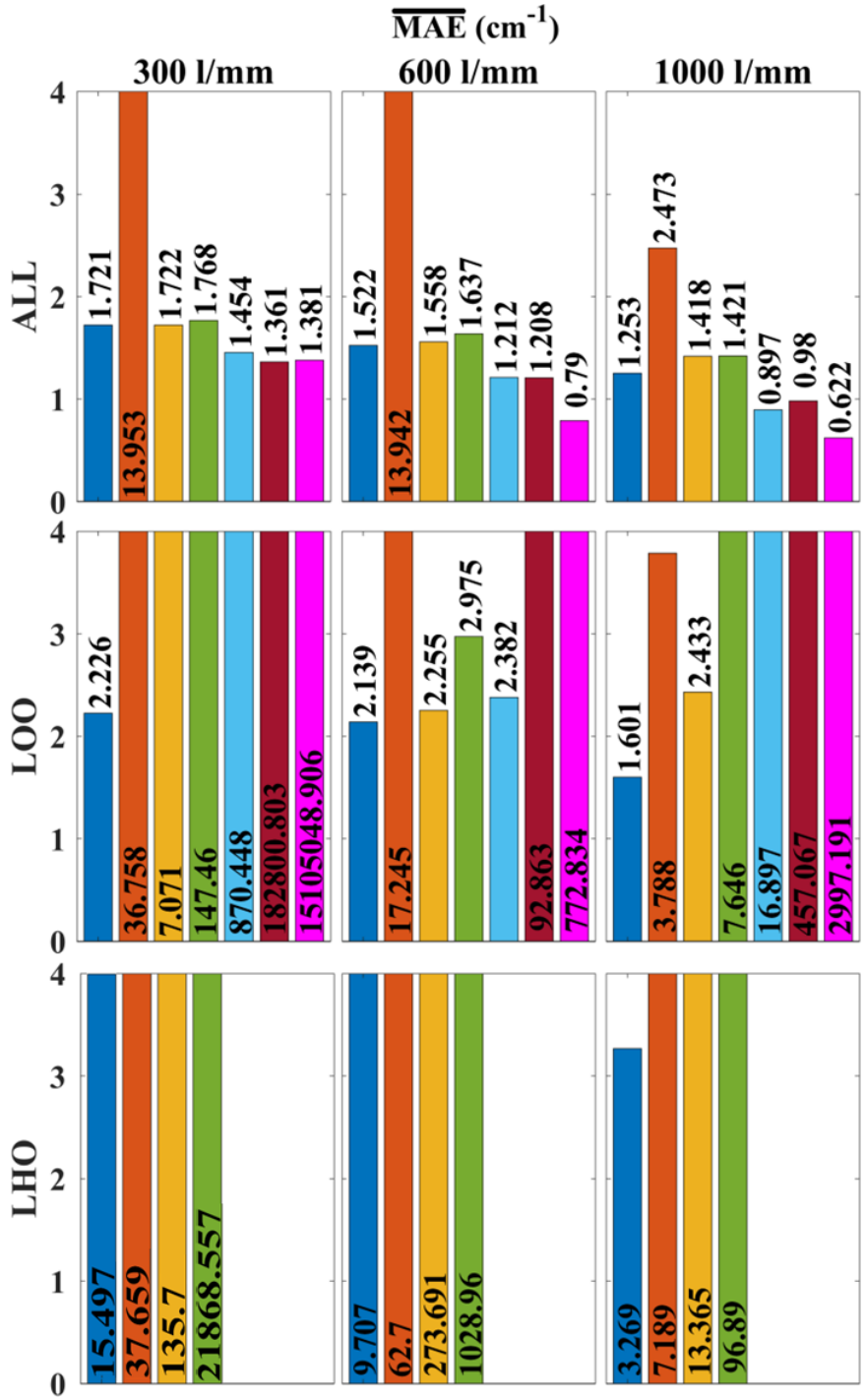


Figure 5.6: Evaluation of direct wavenumber calibration accuracy using Mean Absolute Error applied to commercial polymer spectra. The reference peak position of polymer is based on the result of our database, which was shown in Table 3.1. In this case the transmission spectrometer was not tested. Polynomial order >3 could not be applied for LHO evaluation due to lower peak number. The results of the algorithm proposed in this chapter are given in blue and the results for first-, second-, third, fourth, fifth, and sixth-order polynomial fitting are given in orange, yellow, green, blue, red, and pink, respectively.

Similar trends are reported for the benzonitrile reference spectrum as shown in Fig. 5.5. In this case, it was not possible to perform polynomial fitting with order > 3 for LHO evaluation owing to the availability of a smaller number of lines in the reference spectrum. Focusing only on LOO and LHO evaluation, which provides the best estimate of calibration accuracy inside and outside of the reference lines, it is notable that the proposed algorithm has the best accuracy in all cases, with the most pronounced improvement over polynomial fitting observed for LHO evaluation. For LOO evaluation third-order fitting is the second most accurate in all cases; in summary, the proposed algorithm provides an accuracy of 0.105-0.736 cm^{-1} for the four spectrometers within the bounds of the reference lines, while third-order fitting produces errors that are 1.6-2.7 times greater. For LHO evaluation, second-order fitting is the second most accurate in all cases; in summary, the proposed algorithm provides an accuracy of 0.223-0.967 cm^{-1} for the four spectrometers outside the bounds of the reference lines, while second-order fitting produces errors that are 1.8-9.2 times greater, and third order fitting produces errors that are 23.8-178.8 times greater.

Although the polymer material has the advantage of photo-stability and easy mounting, it produces the least accurate wavenumber calibration results across the three materials tested. Nevertheless, it provides similar trends as for the other two cases as shown in Fig. 5.6. As for benzonitrile, it was not possible to perform polynomial fitting with order > 3 for LHO evaluation. Focusing again only on LOO and LHO evaluation, it is clear that the proposed algorithm has the best accuracy compared with polynomial fitting of various orders, albeit the improvement over second order fitting is minute for the 600 lines/mm grating. As before, the most significant improvement over polynomial fitting is observed for LHO evaluation. For LOO evaluation second-order fitting is the second most accurate in all cases; in summary, the proposed algorithm provides an accuracy of 1.601-2.226 cm^{-1} for the three spectrometers within the bounds of the reference lines, while second-order fitting produces errors that are 1.1-3.2 times greater and third order fitting produces errors that are 1.4-66.2 greater. For LHO evaluation, first-order fitting is the second most accurate in all cases; in summary, the proposed algorithm

provides an accuracy of 3.269-15.497 cm^{-1} for the four spectrometers within the bounds of the reference lines, while first-order fitting produces errors that are 2.2-6.9.0 times greater.

In Appendix 5.8.1 we provide the same evaluations for the three materials, where the underlying metric of mean absolute error is replaced with the standard deviation, and the root mean square error, which are sometimes preferred in the literature.

5.7 Conclusion

In this chapter, we have made several important contributions in the area of direct wavenumber calibration for Raman spectroscopy. The first contribution is the derivation of the relationship between the detector pixel position x and wavenumber υ for a low distortion Raman spectrometer, in terms of the system parameters including, laser wavelength, grating period and angle, and spectrograph focal length, which is provided in Section 5.2.2 and which build on the result in Section 5.6 in the previous chapter. This relationship was explored for a number of different experimental systems, and it was demonstrated that the degree of non-linearity was highly variable across the different systems and depended primarily on focal length and grating period; in some cases a second order fit could estimate the relationship with high accuracy, while in others this would result in high error and a third order or higher polynomial order is necessary for accurate fitting. We believe that this result goes some way to explain the variable results that have been reported in the literature to date on the optimal polynomial order to be used in direct wavenumber calibration.

The most significant contribution in the chapter is the algorithm proposed in Section 5.3, which can replace the polynomial fitting step applied in traditional wavenumber calibration to relate x and υ. This algorithm searches for the optimal set of system parameters as functions of the $x-\upsilon$ expression derived in Section 5.2.2, that can best explain the positions and wavenumber values for a set of spectral lines in a known reference spectrum. The algorithm can search over seven system parameters in total including the laser wavelength and is demonstrated to outperform traditional polynomial fitting in terms of a number of metrics.

In order to fully demonstrate the superiority of the proposed algorithm over polynomial fitting, we employ the metrics developed in the previous chapter in Section 5.5.2, namely leave-one-out and leave-half-out cross validation. Although these are well known techniques in chemometrics, their applicability to wavenumber calibration has not previously been reported. We argue that these matrices are more suitable than traditional approaches as they preclude the possibility of over-fitting in the calibration

process; by only allowing the accuracy to be measured on peaks that were not included in the calibration process, these methods must be considered to be a more accurate evaluation of the true accuracy of the calibration procedure at wavenumber positions between the spectral lines in the reference wave and in bands outside of the outermost peaks in the reference spectrum. Using these evaluations, we conclude that for the instruments tested here, the proposed algorithm is more accurate than second- or third-order fitting within the band of the spectral lines in the reference by a factor of up to 2.2 times for 4-acetamidophenol and 5.9 times for benzonitrile. More significantly it is more accurate than second- or third- order fitting outside of the reference lines by factors of up to 269.6 times and 176.9 times these two materials.

Another interesting conclusion is that benzonitrile provides for more accurate calibration than 4-acetamidophenol for all four spectrometers tested: the accuracy afforded by 4-acetamidophenol is LOO:0.254-0.997 cm^{-1} and LHO:0.295-1.118 cm^{-1} for the four systems, while for benzonitrile this drops to LOO:0.105-0.736 cm^{-1} and LHO: 0.223 - 0.967 cm^{-1}. The latter has significantly fewer peaks; however, these peaks are in general sharper, which may suggest a greater importance for peak width compared with peak number for wavenumber calibration. It is important to emphasize that the proposed algorithm negates the need for a large number of peaks. The accuracy when using only nine peaks for the case of the 1000 lines/mm grating is ALL:0.072 cm^{-1}, LOO:0.105 cm^{-1}, LHO: 0.223 cm^{-1}. These values compare well with the most accurate calibration reported to date, [39] which was limited only by the accuracy of the ASTM values for the reference lines of 0.1 cm^{-1}. In that paper the authors used 67 peaks from a reference spectrum from a composite of different materials, chosen to cover a wide range in wavenumber. The proposed algorithm may negate the need for such an approach, and we believe there is a strong case for it to be included in future iterations of ASTM-E1840, in particular for wavenumber bands outside the range of lines in the reference spectrum.

In this chapter and the previous two chapters the main subject was wavenumber (or wavelength) calibration for Raman spectrometers. In the next chapter we change direction and focus on intensity calibration.

5.8 Appendix

5.8.1 Results using Standard Deviation and RMSE

In this appendix ALL, LOO, and LHO evaluation of the proposed algorithm is repeated for the underlying metrics of standard deviation and root mean square error, which are sometimes preferred in the literature. The definitions of these metrics are given in Section 5.5.2. Evaluation using the error metric of standard deviation is given in Fig. 5.7 ,Fig. 5.9, Fig. 5.11 and using Root Mean Square Error is given in Fig. 5.8, Fig. 5.10, Fig. 5.12. These results correspond to those shown in Fig. 5.4, Fig. 5.5, and Fig. 5.6 in the main body of the chapter for the case of the error metric Mean Absolute Error. These additional results are shown here to help in comparing with results from other papers.

5.8.2 Detailed analysis of peak error

In this section, we examine the wavenumber error for each individual peak across the three reference spectra using each of the different metrics: ALL, LOO, and LHO. The goal is to investigate patterns in the wavenumber error when using the different calibration methods. More explicitly, we calculate the peak mean absolute error (PMAE) for each reference line v_i in the reference spectrum, which is defined as follows:

$$\overline{PMAE(v_i)} = \frac{1}{K}\sum_{k=1}^{K} |error_k(v_i)| \tag{5.21}$$

where $error_k(v_i)$ is given by the function $error(v_i)$ for the k^{th} spectrum as defined in Equation 5.10 in the main text of the chapter, and again K is the number of reference spectra that have been recorded, which is $K = 100$ for the experiments in this chapter. The $PMAE$ function is calculated for all M spectral lines in the reference spectrum. It should be noted that the MAE function defined in Equation 5.18 can be rewritten in terms of this peak error function as follows:

$$\overline{MAE} = \frac{1}{M}\sum_{i=1}^{M} PMAE(v_i) \tag{5.22}$$

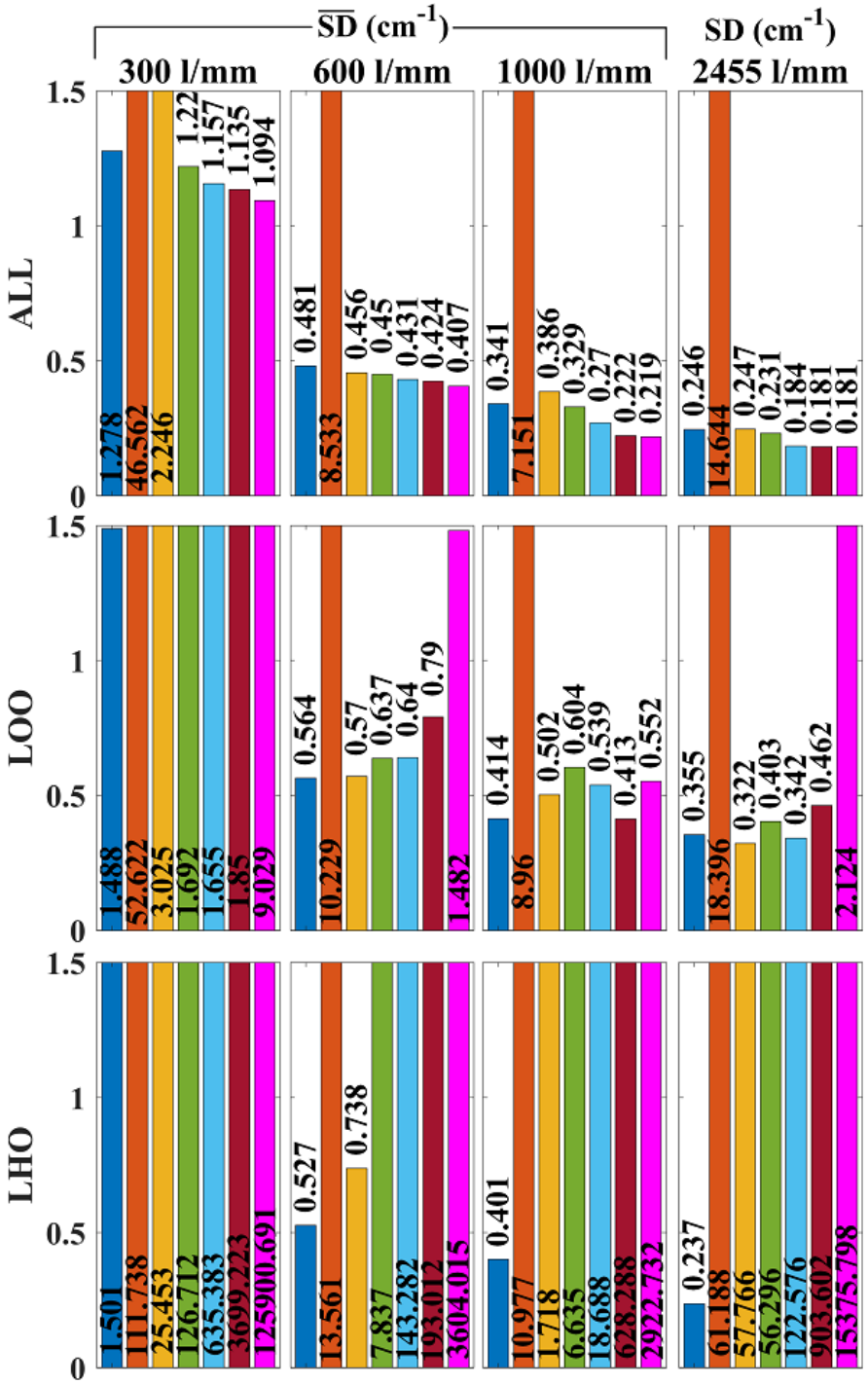


Figure 5.7: Wavenumber errors for 4-acetamidophenol using the standard deviation. The results of the algorithm proposed in this chapter are given in blue and the results for first-, second-, third, fourth, fifth, sixth, and seventh-order polynomial fitting are given in orange, yellow, green, blue, red, and pink, respectively.

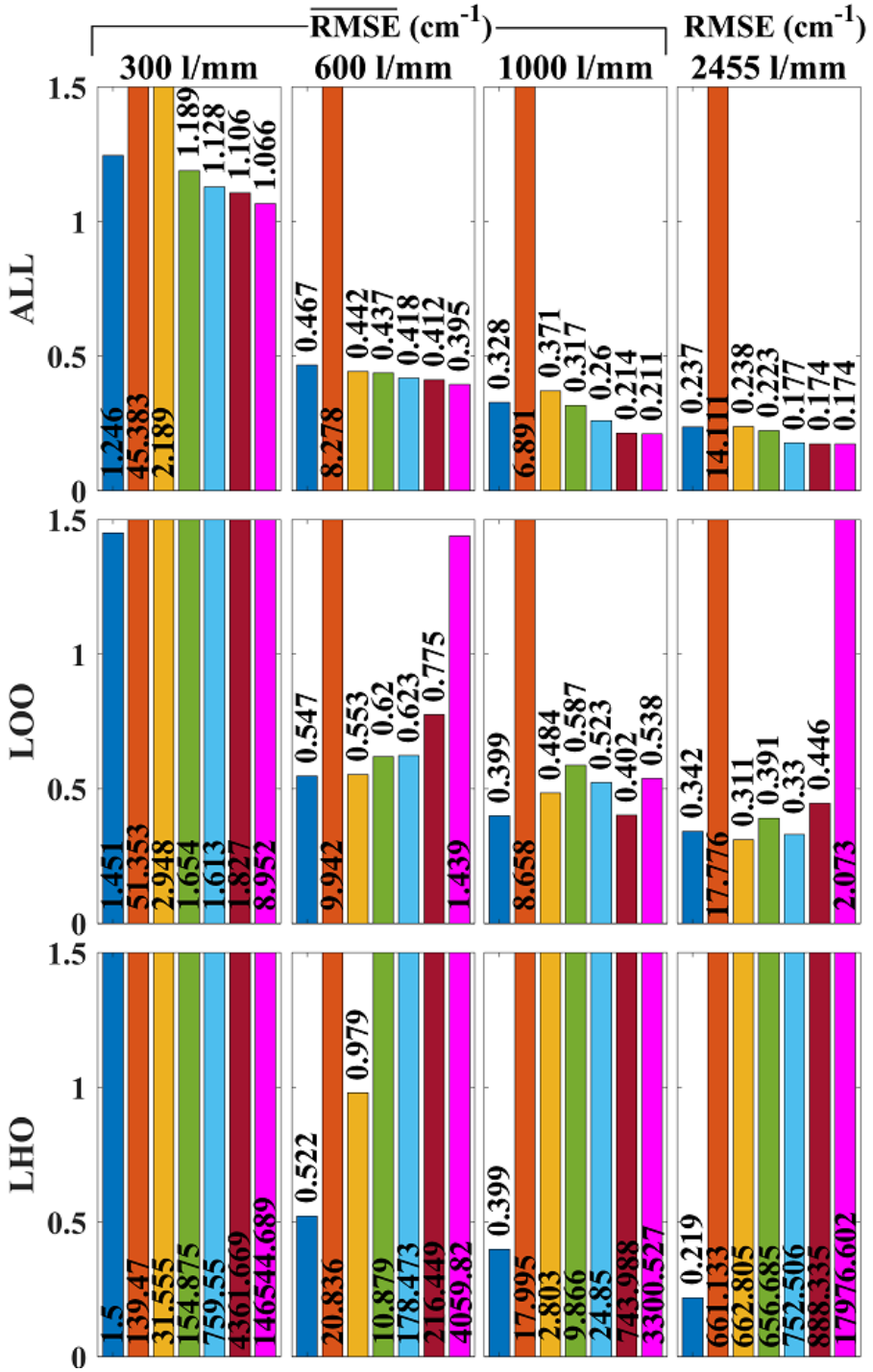


Figure 5.8: Wavenumber errors for 4-acetamidophenol using the RMSE. The results of the algorithm proposed in this chapter are given in blue and the results for first-, second-, third, fourth, fifth, sixth, and seventh-order polynomial fitting are given in orange, yellow, green, blue, red, and pink, respectively.

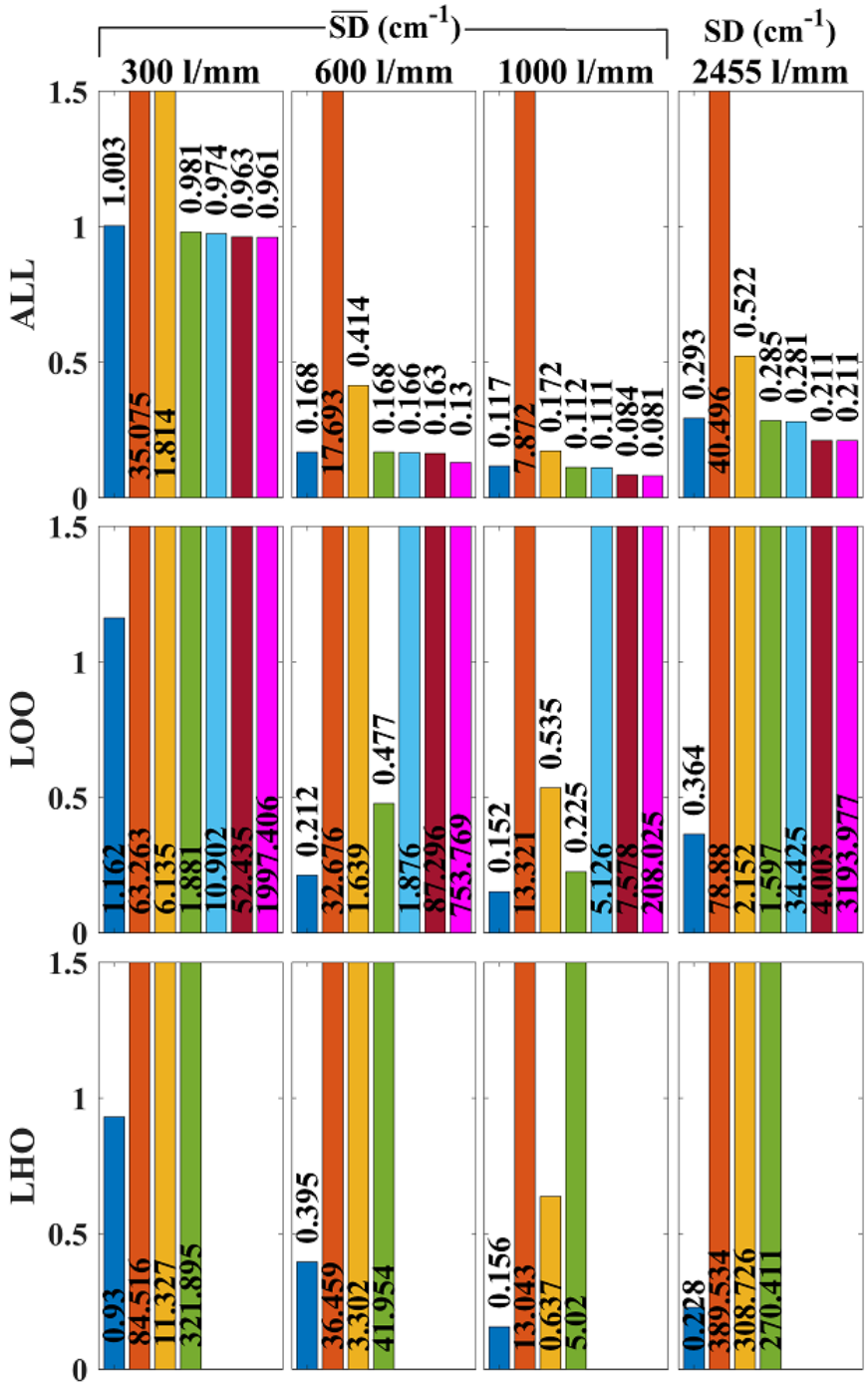


Figure 5.9: Wavenumber errors for benzonitrile using the standard deviation. The results of the algorithm proposed in this chapter are given in blue and the results for first-, second-, third, fourth, fifth, sixth, and seventh-order polynomial fitting are given in orange, yellow, green, blue, red, and pink, respectively.

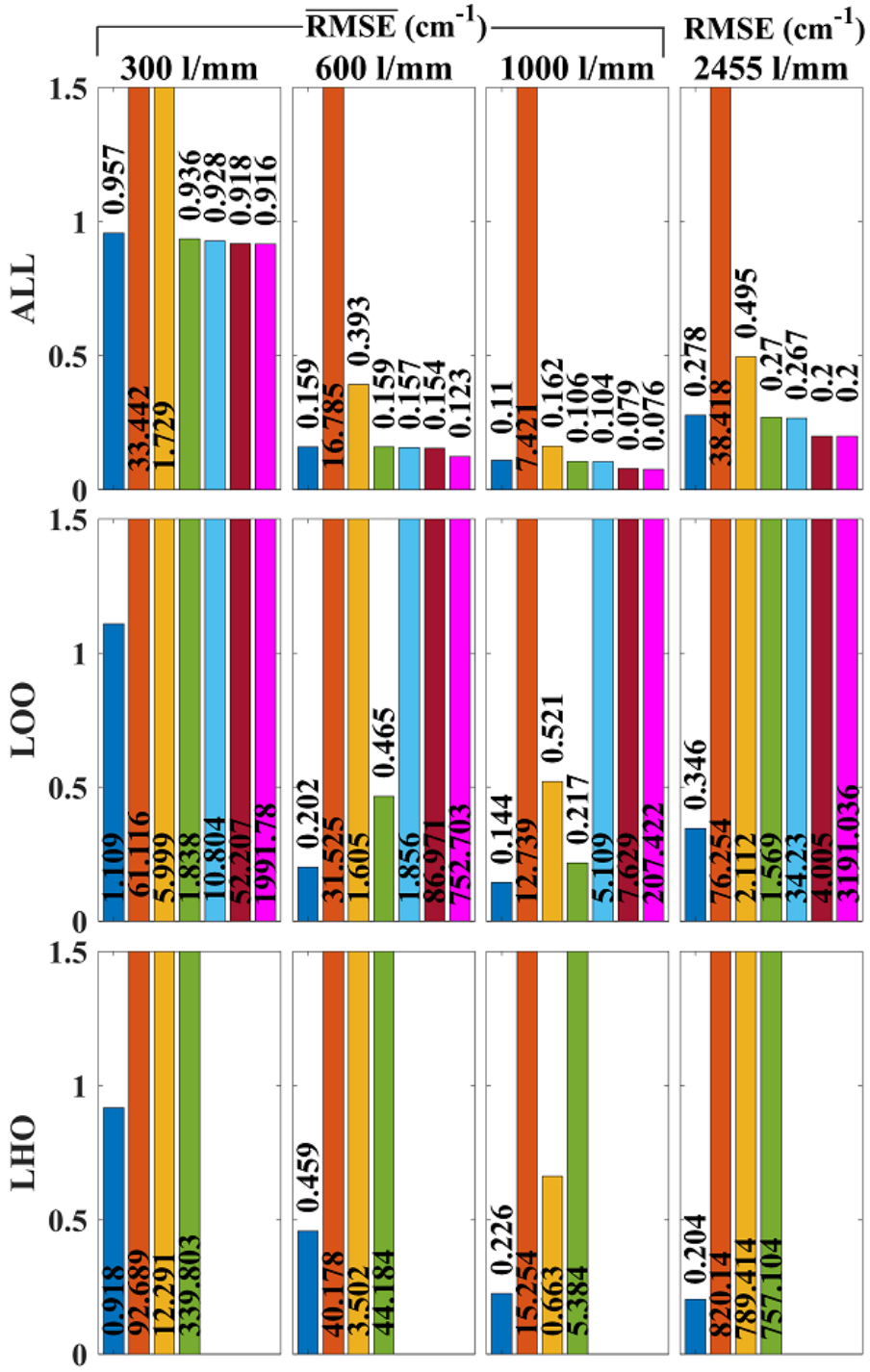


Figure 5.10: Wavenumber errors for benzonitrile using the RMSE. The results of the algorithm proposed in this chapter are given in blue and the results for first-, second-, third, fourth, fifth, sixth, and seventh-order polynomial fitting are given in orange, yellow, green, blue, red, and pink, respectively.

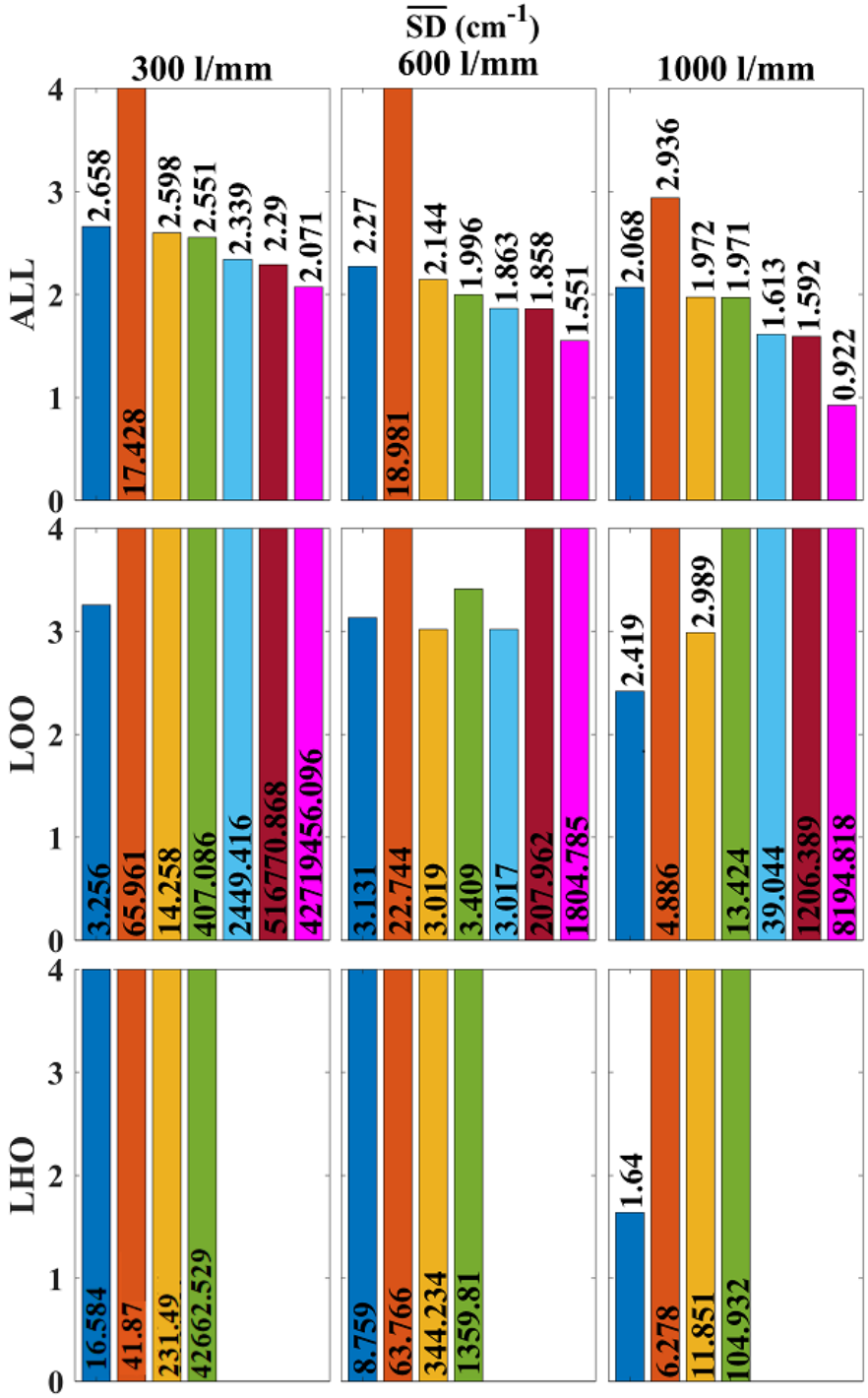


Figure 5.11: Wavenumber errors for commercial polymer using the standard deviation. The results of the algorithm proposed in this chapter are given in blue and the results for first-, second-, third, fourth, fifth, sixth, and seventh-order polynomial fitting are given in orange, yellow, green, blue, red, and pink, respectively.

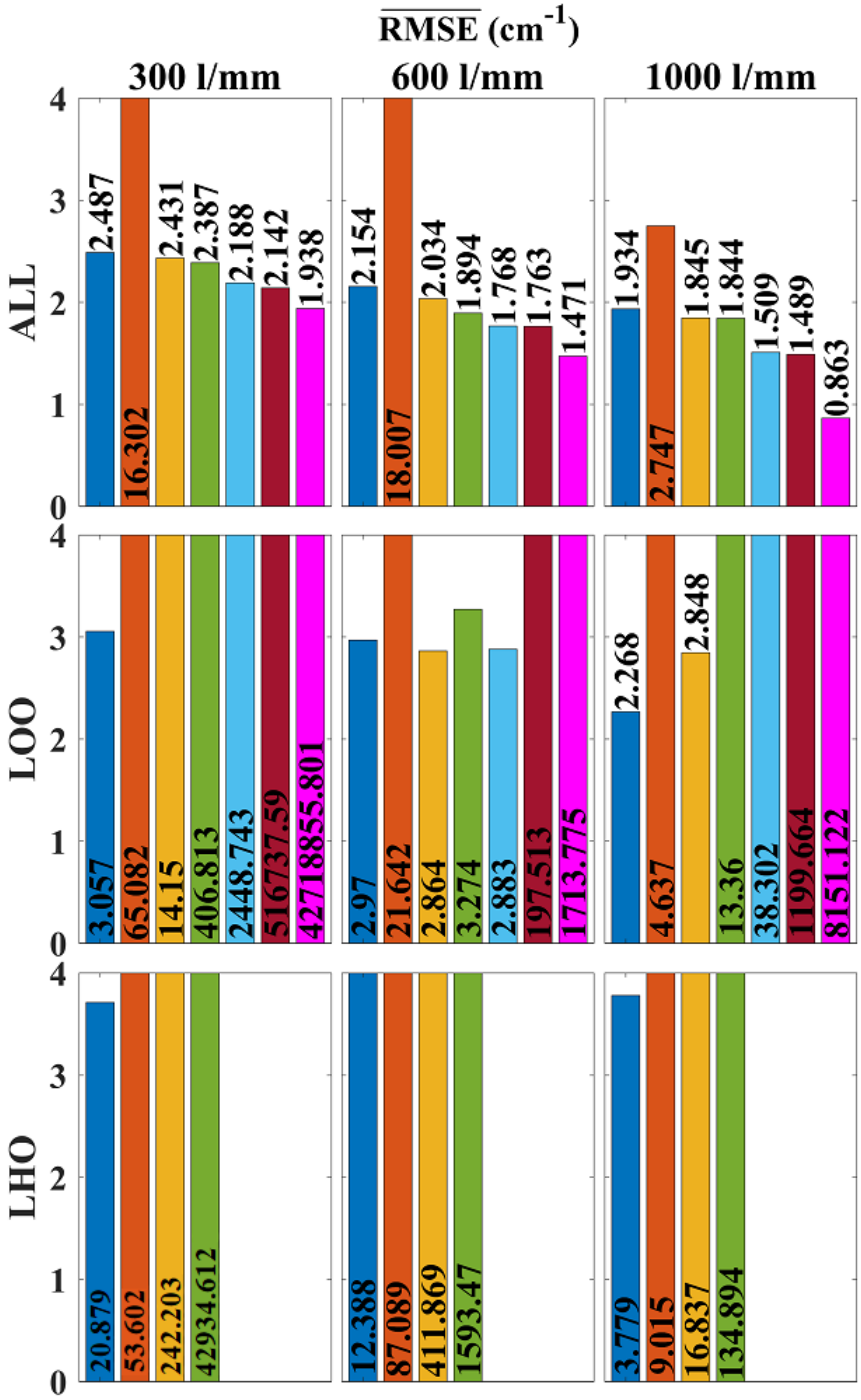


Figure 5.12: Wavenumber errors for commercial polymer using the RMSE. The results of the algorithm proposed in this chapter are given in blue and the results for first-, second-, third, fourth, fifth, sixth, and seventh-order polynomial fitting are given in orange, yellow, green, blue, red, and pink, respectively.

We begin in Fig. 5.13 by calculating the PMAE function for the wavenumber reference 4-acetamidophenol spectra that have been wavenumber calibrated using 2nd order polynomial fitting, 3rd order polynomial fitting, and using the method proposed in this chapter. It should be noted that these error functions have been calculated over a dataset of $K = 100$ different reference spectra that have been recorded with movements of the grating angle. In Fig. 5.13 (a1) the PMAE is shown for the 20 reference peaks in 4-acetamidophenol for the case of using ALL peaks and for the case of using LOO analysis. For ALL peaks, all 20 peaks were used in the calibration routine, and then the $error_k(\upsilon_i)$ function for each υ_i was measured for the calibrated axis. Therefore, for ALL peaks the calibration routine is applied once for each of the $K = 100$ spectra. As described in the chapter, for LOO the $error_k(\upsilon_i)$ function is calculated on a reference line υ_i, that has not been included in the calibration protocol; therefore, for a given spectrum, k, each calibration protocol is applied 20 times, using a different set of 19 peaks in each instance, and the missing wavenumber υ_i is used to calculate the peak error defined by Equation 5.10. This is repeated for each of the 100 spectra in the dataset to obtain the PMAE as defined by Equation 5.21; therefore, each calibration routine is applied 2000 times for LOO analysis. Results are shown for 2nd and 3rd order fitting as well as for the proposed calibration method in Fig. 5.13 (a1). For each of the three calibration protocols it can be seen that LOO is slightly less accurate that the ALL peaks case, which is to be expected since LOO removes the possibility of overfitting and provides a more reliable evaluation of how a calibration protocol will perform at wavenumber values that are not included in the discrete set of reference wavenumber lines. It is interesting to note the similar pattern of relative error across the different lines in the spectrum, particularly for 3rd order fitting and for the proposed algorithm. It is also interesting to note that for both calibration protocols the highest errors occur for the peaks at 213 cm^{-1} and 1515 cm^{-1}, which are amongst the broadest and weakest lines in the spectrum; sharper peaks have lower error, which may point to a difficulty in accurately determining sub-pixel position for broader peaks in general. Notably, the peak at 213 cm^{-1} also has one of the highest uncertainties of all of the peaks (see

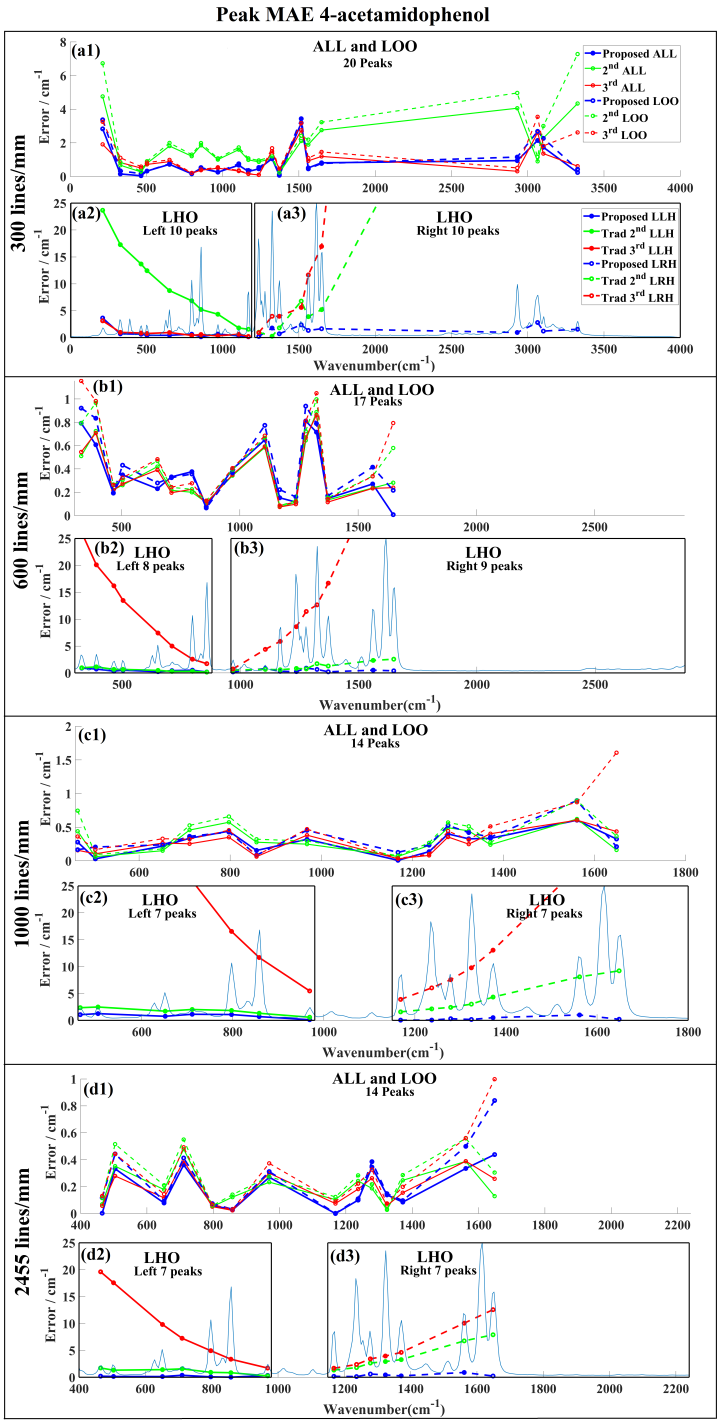


Figure 5.13: Peak Mean Absolute Error (PMAE) calculated for 4-acetamidophenol spectra that have been wavenumber calibrated using 2nd order polynomial fitting, 3rd order polynomial fitting, and using the method proposed in this chapter. It should be noted that these error functions have been calculated over a dataset of 100 different reference spectra that have been recorded with movements of the grating angle. Results for ALL, LOO and LHO analysis are shown in (a1-a3) for the 300 lines/mm grating; in (b1-b3) for the 600 lines/mm grating; in (c1-c3) for the 1000 lines/mm grating; and in (d1-d3) for the 2455 lines/mm grating. For the latter case the dataset contains only a single spectrum since the grating could not be rotated. See text for more details.

Table 5.3). Inspection of the standard deviation values in the table reveals that those peaks with the smallest deviations appear to have the smallest errors in the figure. In Fig. 5.13 (a2) and (a3) the results of LHO analysis are presented. Fig. 5.13 (a2) shows the error for the left 10 peaks when calibration has been applied using only the right 10 peaks. Surprisingly, third order fitting performs well in this instance; however, this appears to be a fortuitous result and when we tested again using the right 9 most peaks for calibration significant error appeared for the left 11 peaks. In Fig. 5.13 (a3) the error for the right most 10 peaks is shown when calibration used only the left 10 peaks; in this case only the proposed method performs well. The disparity in error in Fig. 5.13 (a2) and (a3) relates to the range and distribution of the 10 peaks used for calibration. The same set of results are shown for the 600 lines/mm grating in Fig. 5.13 (b1-b3). Once again the proposed method is the most consistent overall. Interestingly, in this case 3rd order fitting performs poorly for LHO analysis for both sides, while 2nd order fitting performs better. Similar results are shown for the 1000 lines/mm grating in Fig. 5.13 (c1-c3) and for the 2455 grating in Fig. 5.13 (d1-d3); in the later case only a single spectrum is used in the dataset, i.e. $K = 1$. It is notable that the same approximate pattern of error is found for the different gratings in terms of the relative error of each peak in the 4-acetamidophenol spectrum.

A similar set of results are shown in Fig. 5.14 for the benzonitrile reference material for the 300 lines/mm grating in (a1-a3), 600 lines/mm grating (b1-b3), 1000 lines/mm grating in (c1-c3) and the 2455 lines/mm grating in (d1-d3). As for the previous case LOO is slightly worse than ALL for all methods and the proposed method shows slightly better error performance when compared with third order fitting for both ALL and LOO for all four gratings. However, the proposed method is clearly superior for LHO analysis is all cases.

A third set of results are shown in Fig. 5.15 for the polymer reference material for the 300 lines/mm grating in (a1-a3), 600 lines/mm grating (b1-b3), and the 1000 lines/mm grating in (c1-c3). For the case of the 300 lines/mm grating not all of the lines in Table 5.3 could be used due to the relatively low resolution of that system causing neighbouring

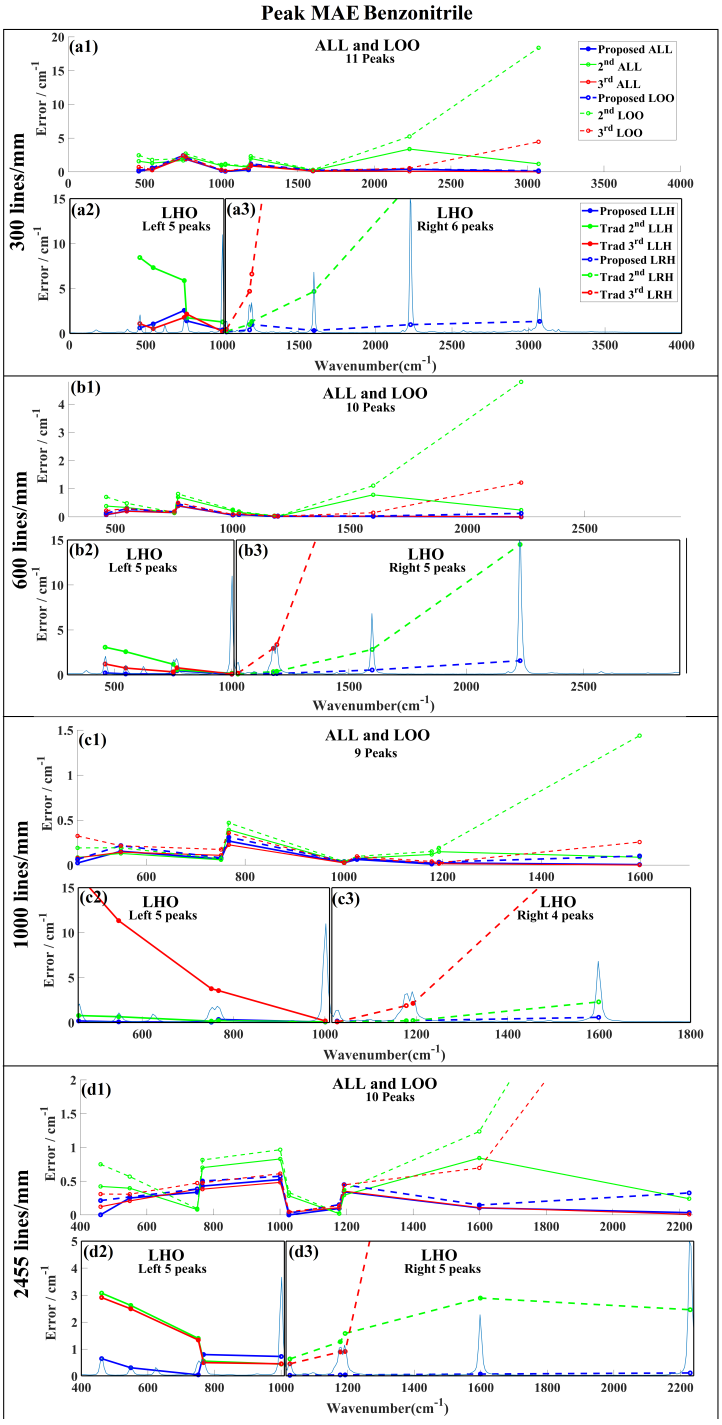


Figure 5.14: Peak Mean Absolute Error (PMAE) calculated for benzonitrile spectra that have been wavenumber calibrated using 2nd order polynomial fitting, 3rd order polynomial fitting, and using the method proposed in this chapter. It should be noted that these error functions have been calculated over a dataset of 100 different reference spectra that have been recorded with movements of the grating angle. Results for ALL, LOO and LHO analysis are shown in (a1-a3) for the 300 lines/mm grating; in (b1-b3) for the 600 lines/mm grating; in (c1-c3) for the 1000 lines/mm grating; and in (d1-d3) for the 2455 lines/mm grating. For the latter case the dataset contains only a single spectrum since the grating could not be rotated. See text for more details.

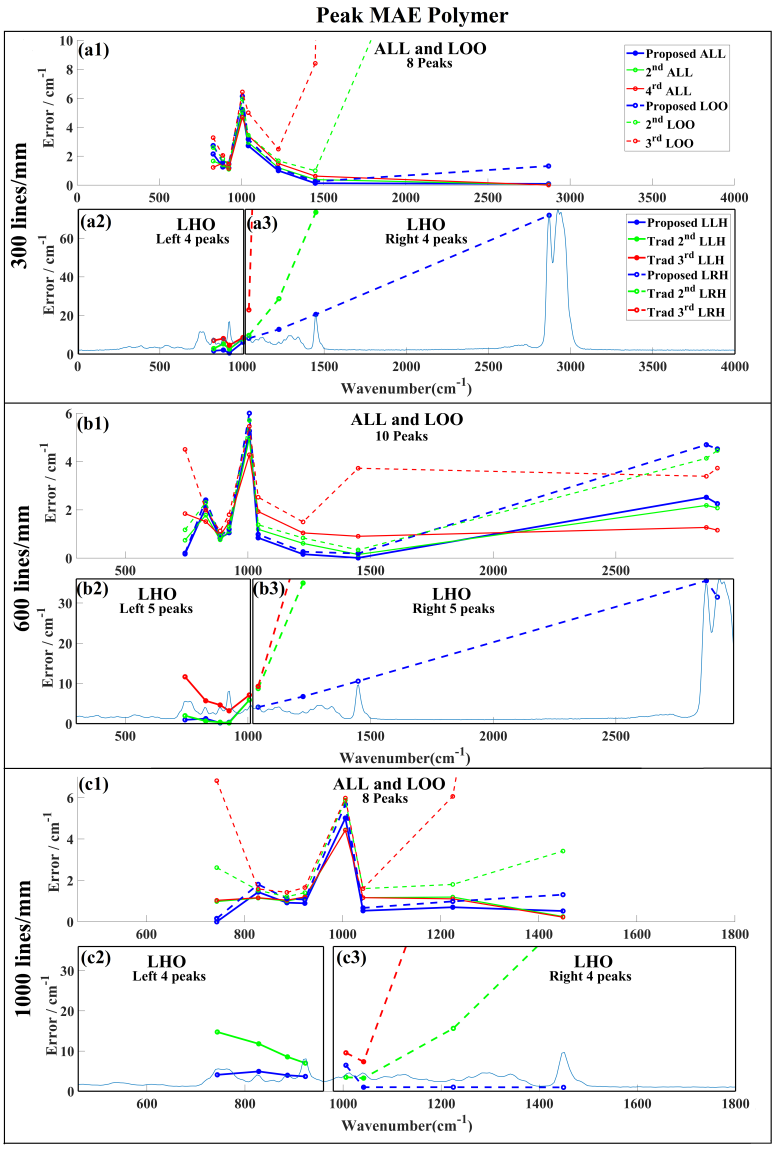


Figure 5.15: Peak Mean Absolute Error (PMAE) calculated for the polymer spectra that have been wavenumber calibrated using 2nd order polynomial fitting, 3rd order polynomial fitting, and using the method proposed in this chapter. It should be noted that these error functions have been calculated over a dataset of 100 different reference spectra that have been recorded with movements of the grating angle. Results for ALL, LOO and LHO analysis are shown in (a1-a3) for the 300 lines/mm grating; in (b1-b3) for the 600 lines/mm grating; in (c1-c3) for the 1000 lines/mm grating; See text for more details.

peaks to blur and merge. For this reason, the two lines at 743 cm^{-1} and 2914 cm^{-1} were omitted. For the other two gratings all peaks listed in the table that were within the bandwidth of the spectrometer were included in the calibration and error analysis. There are a number of interesting points in relation to the results. Firstly, the peak at 1005.7 cm^{-1} consistently has the worst for each of the ALL and LOO analyses in Fig. 5.15 (a1), (b1) and (c1); this is in spite of the fact that this peak is reported to have the lowest uncertainty of all of the peaks (see Table 5.3); this indicates that there may be a problem with the values reported by Ref. 3. The second point of interest is the high error for the right most peak at 2869 cm^{-1} in the LOO analysis in Fig. 5.15 (a1) and the high error of two right most peaks at 2869 cm^{-1} and 2914 cm^{-1} in the LOO analysis in Fig. 5.15 (b1). These high errors result from the long distance between these right most peaks and the other peaks, which appear in the fingerprint region. Such a long distance clearly has a deleterious effect on the results for 2nd and 3rd order fitting.

Chapter 6

Intensity Calibration of Raman Spectrometer using Arbitrary White Light

The work in this chapter has been prepared as a journal paper and will be submitted to the Journal of Raman Spectroscopy with the following reference: *Liu, Dongyue, and Bryan M. Hennelly. "Intensity Calibration of Raman Spectrometer using Arbitrary White Light." to be submitted to Journal of Raman Spectroscopy Nov 2022* with the following abstract:

"In this paper, we develop a simple protocol that enables the use of any uncalibrated white light source for Raman intensity calibration. The basis for this approach is (i) the assumption that the true white light spectrum can be modeled by a low order polynomial and (ii) the use of an easily accessible and high purity reference Raman sample such as glycerol. Unfortunately, the reference glycerol spectrum cannot be used as a single standard reference material due to its low intensity in large bands. However, when combined with the arbitrary white light the correction factor can be estimated. Our results demonstrate that this method performs better than traditional known white light calibration in almost all cases tested."

6.1 Introduction

In all of the previous contributory chapters, i.e. Chapters 3, 4, and 5, the core subject was wavenumber (or wavelength) calibration. In this chapter we change direction and focus instead on intensity calibration for a given Raman spectrometer. Often, intensity calibration will require a wavelength calibration protocol, such as that described in Chapter 4, together with a carefully calibrated white-light lamp. In this chapter we aim to remove the requirement for the white-light lamp to be be calibrated in advance. A key component in this approach is the use of a Raman material standard. The polymer investigated in Chapter 3 is now reinvented as an intensity calibration tool, perhaps providing a dual purpose.

Raman scattering occurs when an incident laser photon interacts with a molecular bond and excites a vibration/rotational mode. The consequence of this excitation is the emission of scattered photons that have lesser energy than the laser photons; this translates to an increase in wavelength. Since different materials will result in the scattering of different patterns of wavelengths, it is often possible to identify the material using the spectrum of light that is emitted from the material following irradiance with a single mode laser. An exciting area of research is Raman diagnostics, [12–14] whereby the spectrum from a diseased tissue or cell sample can be distinguished from that of a healthy sample making use of multivariate statistical methods [12–14, 17–22]. Recently, there have been a number of significant steps in moving Raman one step closer to the clinic, including Raman guided surgery [30–32] and automated Raman cytology. [33–37]

However, despite the huge promise of RS as a non-invasive and ubiquitous diagnostic platform, clinical progress has been severely hindered by poor cross-instrument comparability. Two recent multi-site studies, [44, 45] have highlighted inconsistencies in the spectra recorded from different materials across several different instruments. Cross-instrument differences relate to both wavenumber shift as well as intensity variation for the same sample, and it is the latter problem that this chapter aims to address. Intensity variation across instruments is caused by the differing wavelength and polarisation dependent transmission function of each instrument and possibly differing

instrument resolution. The difference in intensity response from two instruments is most obvious when recording the Raman spectrum from the same sample using two different excitation wavelengths such as 532nm and 785nm.

There has already been some development of consensus standards for Raman instrumentation by the American Society for Testing and Materials (ASTM) International in relation to performance testing, calibration, and relative intensity correction (ASTM E1683 [46], E1840 [2], E2529 [47], E2911 [48]). An excellent review of these standards is provided in Refs 49 and 40. In the context of this chapter, the most relevant of these standards documents is ASTM-E2911, [48] most recently updated in 2013, which focuses on Raman intensity calibration using the NIST Standard Reference Materials (SRMs) in the 224X series. These luminescent glass materials have been calibrated at NIST for a variety of particular excitation wavelengths and have high photostablity. The glass is placed in the Raman spectrometer and illuminated with the laser. The resulting fluorescent spectrum is recorded and divided by the known spectrum from NIST in order to obtain the correction factor to be applied to all Raman spectra recorded by the same instrument.

Despite the fact that Ref. 48 is the only published standards document on intensity calibration, the most commonly applied method of Raman intensity calibration in the literature is white-light calibration, whereby a tungsten-halogen with known spectrum (measured using a reference that can be traced back to NIST or a similar organisation) is recorded using the spectrometer. Similar to SRMs, the recorded spectrum is divided by the known spectrum in order to obtain the correction factor to be applied to all Raman spectra recorded by the same instrument. Perhaps the main reason why white-light calibration is preferred over fluorescence based SRMs is the ubiquity of the approach. An SRM must be purchased for a specific excitation wavelength while the calibrated white light source can be applied to spectrometers with all excitation wavelengths. In addition, the cost of a single SRM is higher than a white-light lamp. Another reason may be that SRMs have a positional dependency; the recorded spectrum can change appreciably if the SRM is not positioned reliably at the laser focus and moves along the

optical axis. [48]

White light calibration lamps can also be problematic. Although relatively less expensive than SRMs, they are still high cost and must be regularly re-calibrated after approximately 50 hours. This re-calibration is also costly and time consuming and it also begs the question on how accurately the lamp spectrum is known after 50 hours. Another issue that commonly presents is the positioning of the lamp in the setup. Often a fiber is used to deliver the light but also can present mounting problems.

In this chapter, we develop a simple protocol that enables the use of any uncalibrated white light source for Raman intensity calibration. The basis for this approach is (i) the assumption that the true white light spectrum can be modeled by a low order polynomial and (ii) the use of an easily accessible and high purity reference Raman sample such as glycerol. Unfortunately the reference glycerol spectrum cannot be used as a single SRM material due to its low intensity in large bands. However, when combined with the arbitrary white light the correction factor can be estimated. Our results demonstrate that this method performs better than traditional known white light calibration in almost all cases tested.

6.2 Spectral Irradiance of an Incandescent Lamp

Tungsten-halogen incandescent lamps are thermal radiators based on the generation of light of various wavelengths by heating of the filament to a very high temperature. In contrast to black bodies, which are ideally predictable radiators with known emissivity and spectral constancy, incandescent lamps are not perfect Planckian radiators; more specifically, the spectral behaviour of Tungsten-halogen lamps varies as a function of the emissivity, which in turn varies as a function of temperature. [141] The spectral irradiance, $E(\lambda, T)$, of a Tungsten-Halogen incandescent lamp is modelled by the following equation:

$$E(\lambda, T) = B(T)\epsilon_W(\lambda, T)\epsilon_{\triangle}(\lambda)\frac{2hc^2}{\lambda^5\left[\exp\left(\frac{hc}{\lambda kT}\right)-1\right]} \tag{6.1}$$

where λ is the wavelength in vacuum, T is the temperature of the filament, $\epsilon_W(\lambda, T)$ is the emissivity of tungsten, $\epsilon_{\triangle}(\lambda)$ is the residual correction factor for the emissivity of the lamp, h is the Planck constant, c is the speed of light in vacuum, and k is the Boltzmann constant. The parameter $B(T)$ is a temperature dependent geometrical factor that captures the effect of the distance from the lamp and the dimensions of the filament. This factor varies as a function of temperature due to thermal expansion. Stable white lights that are typically used for calibration will employ a constant current or constant power source in order to reduce variability. $B(T)$ is defined as follows:

$$B(T) = B_0[1 + 0.000016K_{-1}(T - T_0)] \tag{6.2}$$

The parameter $\epsilon_{\triangle}(\lambda)$, accounts for all additional effects due to various factors such as transmittance of the bulb or filling gas. For the case of a quartz bulb this factor will be smooth over wavelength and will only affect the irradiance by as little as ±2. In Fig. 6.1 we illustrate the irradiance of a Tungsten-Halogen lamp for five values of temperature. Here, we have selected the paramatars: T_0=2770K; B_0=0.6; $\epsilon_W(\lambda, T)$=0.35; and $\epsilon_{\triangle}(\lambda)$=1 which are similar to the simulations provided in Ref. 141.

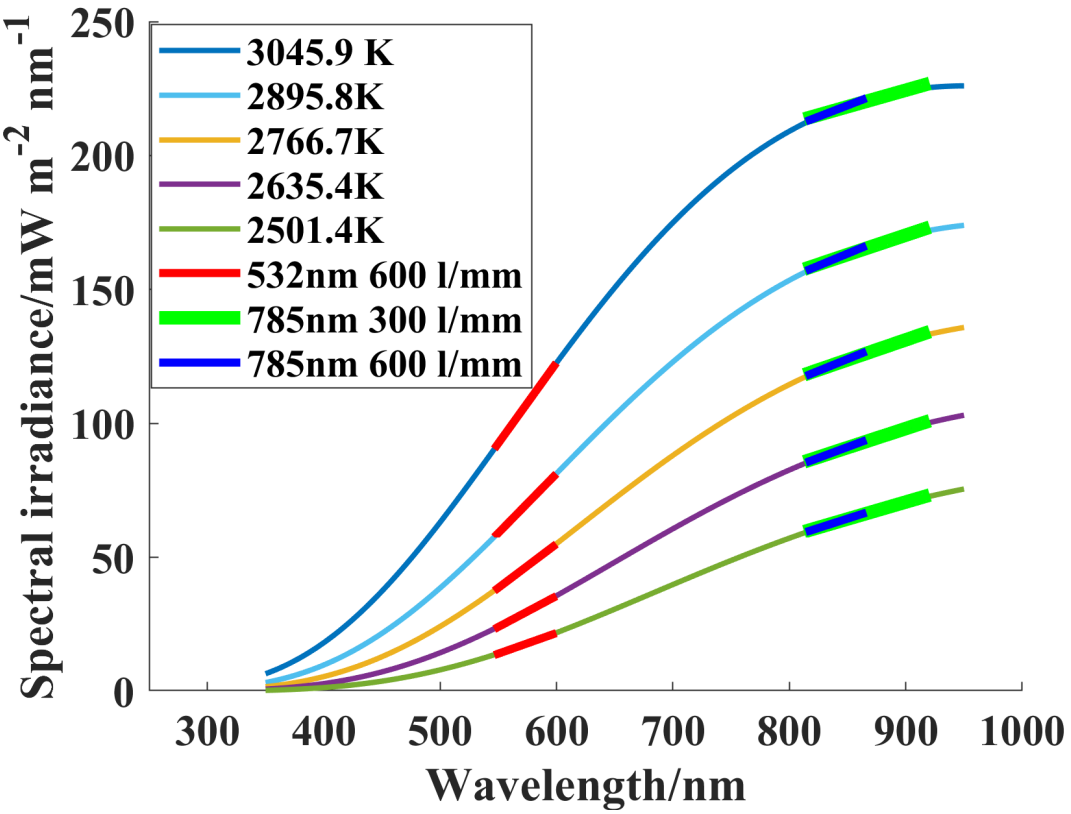


Figure 6.1: Spectral irradiance of a Tungsten Halogen Lamp modelled for different temperatures/currents. Also shown in the figure in thicker blue, black and red lines, are straight line fits to these irradiance profiles in the wavelength bands that correspond to those recorded by the three Raman spectrometers used in this study. The mean absolute error for these fits are provided in Table 6.1

.

T/ K	3045.9	2895.8	2766.7	2635.4	2501.4
	532 nm 600 l/mm				
1st	0.02769	0.02352	0.04513	0.05283	0.05039
2nd	0.00885	0.00670	0.00488	0.00324	0.00189
3rd	0.00017	0.00007	0.00001	0.00002	0.00004
	785nm 300 l/mm				
1st	0.61646	0.48175	0.37194	0.27082	0.18225
2nd	0.01672	0.00894	0.00384	0.00048	0.00230
3rd	0.00034	0.00047	0.0005	0.00047	0.00040
	785nm 600 l/mm				
1st	0.15426	0.11762	0.08863	0.06264	0.04050
2nd	0.0017	0.00077	0.00019	0.00021	0.00044
3rd	0.00003	0.00004	0.00003	0.00003	0.00002

Table 6.1: Mean absolute error for first, second, and third order polynomial fitting applied to the five temperature dependent lamp irradiance profiles shown in Fig. 6.1 in the bands of the three Raman spectrometers used in this study. Units are $mWm^{-2}nm^{-1}$.

An important assumption in this chapter, and one that we seek to validate here, is that the irradiance of any Tungsten-Halogen lamp with unknown emissivity or temperature can be effectively modeled by a low order polynomial over the bandwidth of interest for a typical Raman spectrometer. For the three spectrometers that were used in this study (described in earlier chapters), a straight line fit it applied in the relevant wavelength bands for all five filament temperature values as shown in Fig. 6.1. In order to examine the accuracy of this fit in more detail, these same bands are highlighted in higher resolution in Fig. 6.9. Here, polynomials of order 1, 2, and 3 are fitted to the irradiance values and the corresponding mean absolute error values for these fits are provided in Table 6.1. It is clear that in most cases, a straight line is sufficient to model the irradiance across the spectrometer with low error. Higher accuracy is given by a second order polynomial, while there is no appreciable improvement by using a polynomial order of three or more.

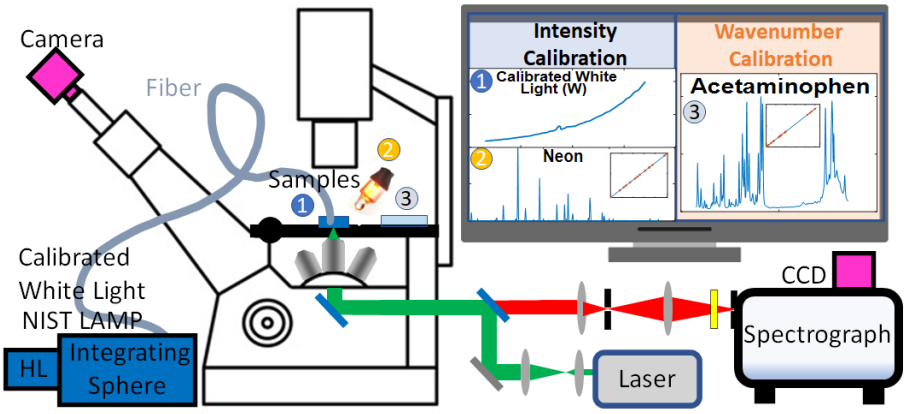
6.3 Traditional Calibration using known White-Light

The classical method for calibrating a Raman spectrum is illustrated in Fig. 6.2; the Raman spectrometer is illustrated in Fig. 6.2a as well as the various samples and lamps commonly used for both intensity and wavenumber calibration. For the former, the first sample of interest is a calibrated white light source. These sources employ a Tungsten-Halogen lamp with controlled power supply and are typically fed into an optical fiber via an integrating sphere or a cosine corrector both of which use diffuse surfaces to couple the light into the fiber. The spectrum of the light that is output from this fiber must be known in advance of performing calibration of the Raman spectrum. This involves frequent calibration of the output every ≈ 50-100 hours, usually performed by a commercial entity, which can be costly and time-consuming. Also, there is the possibility that the lamp can become miscalibrated unexpectedly. The recorded spectrum from the fiber, W, is compared with the known spectrum, W'' in order to obtain a correction factor for intensity calibration of Raman spectra. The second sample of interest is that from an atomic emission source such as neon or krypton, recorded for wavelength calibration in Chapter 3. This is necessitated in order to accurately compare W and W'', which may both have been recorded over a different wavelength axis. The third sample of interest is a wavenumber reference standard such as 4-acetamidophenol or benzonitrile. [2]. The Raman spectrum from this sample is recorded and used to calibrate the wavenumber axis. This third sample is not required for intensity calibration. The various steps involved in recording and processing these spectra for both intensity calibration (Steps 1-3) and wavenumber calibration (Step 4) are detailed in the following subsections.

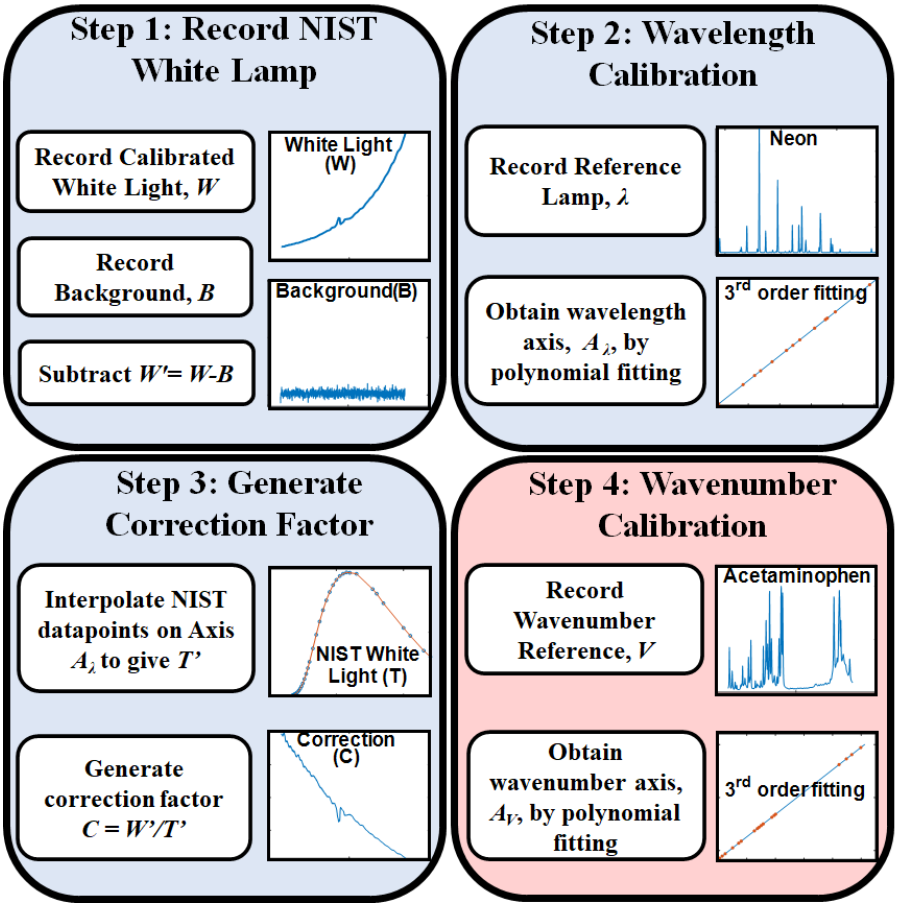
6.3.1 Classical Intensity Calibration

The experimental steps taken for the case of traditional white-light calibration are illustrated in Fig. 6.2b in Steps 1-3 and are described below. We note that all subtraction, multiplication, and division operations are element-wise.

1. The first step is to record the raw spectrum, W, from the calibrated white light



(a) Classical method of intensity and wavenumber calibration in Raman spectroscopy. Spectra from two samples are required for intensity calibration: a calibrated white light source and an atomic emission sample such as neon. Independently, for wavenumber calibration a Raman reference standard is used such as 4-acetamidophenol



(b) Flowchart of Traditional Intensity calibration. In total there are three steps including wavelength calibration. The fourth step, wavenumber calibration, is not necessary but is typically applied as part of the overall calibration process. It is included here to facilitate comparison with the proposed new method described in Fig. 6.3b. See text for further details.

Figure 6.2: Traditional calibration flowchart

source, and subsequently to record the background spectrum, B, of equivalent duration. These are subtracted to remove the influence of the background such that $W' = W - B$. Here W' is a vector of intensity values of length N, where N is the number of pixels in the detector.

2. The second step is to record a spectrum, λ, from an atomic emission lamp such as neon which contains a number of prominent spectral lines over the wavelength bandwidth of the spectrometer. Wavelength calibration is performed on this spectrum as described in Chapter 3 in order to determine the wavelength axis, A_λ. Typically, this involves fitting a third-order polynomial to the pixel position of the spectral lines and their known wavelengths. Following this step, the wavelength axis for the recorded white light spectrum is known and the latter can now be denoted as $[A_\lambda, W']$ where A_λ and W' are both vectors of length N.
3. The third, and final step for intensity calibration is to obtain the correction factor to be applied to all subsequent Raman spectra recorded using this system. The specifications of the calibrated white lamp will include a set of intensity values, T, corresponding to a set of wavelength values, A_T, both vectors of length M. Interpolation is applied to $[A_T, T]$ in order to obtain $[A_\lambda, T']$, where T' is a vector of length N, which represents the true white light intensity values over the wavelength axis A_λ. The correction factor is given by $C = W'/T'$

6.3.2 Wavenumber Calibration

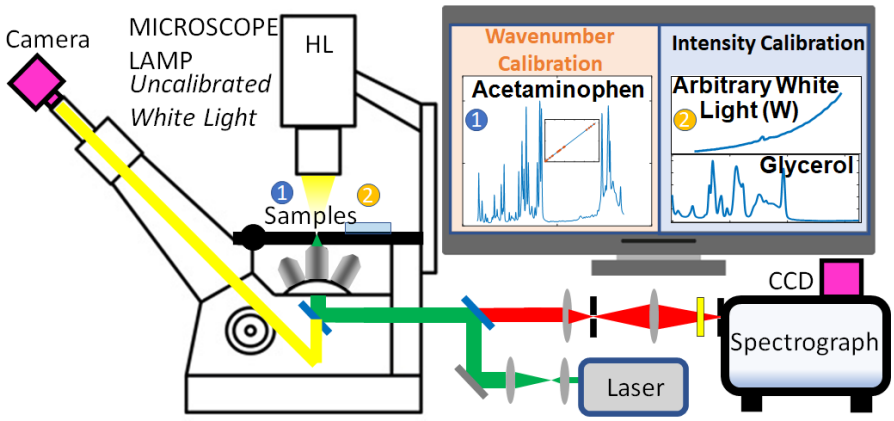
In addition to intensity calibration, it is also necessary to perform wavenumber calibration in order to fully characterise a Raman spectra recorded from the system. This can be achieved by applying wavenumber conversion to the neon spectrum λ. However, this requires measurement of the laser wavelength which can be problematic as discussed in Chapter 4. More commonly, direct wavenumber conversion is applied making use of a Raman wavenumber standard with known spectral lines [2, 142] as described in Chapter 4 and illustrated in Fig. 6.2b Step 4.

4. The spectrum, V from a known wavenumber reference standard such as benzonitrile or 4-acetamidophenol is recorded. It is essential that the chosen standard

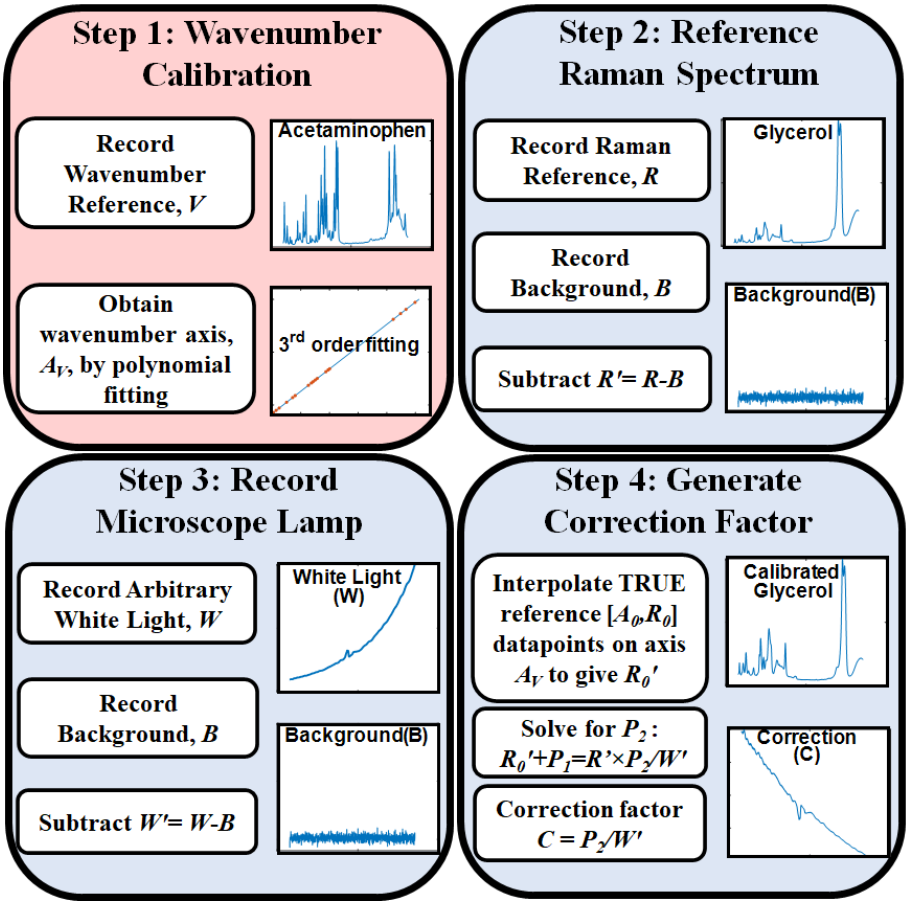
contains spectral lines that span the full width of the wavenumber range recorded by the spectrometer. Several authors have used amalgamated samples in order to achieve this [39] although we have recently demonstrated that high accuracy is possible across the full spectrum even using a small number of peaks as described in Chapter 4. Typically, wavenumber calibration involves third order polynomial fitting applied to the pixel position of the spectral lines in V and their corresponding known wavenumber. [2] Following this step, the wavenumber axis, A_V, is known for any Raman spectrum that is subsequently recorded, where A_V is a vector of length N. This step must be applied frequently as miscalibration can occur for a variety of reasons as described in Section 2.3 in Chapter 2.

6.3.3 Correcting a Raman Spectrum

Correction of a raw sample spectrum, S, makes use of the correction factor C for intensity calibration and the application of the wavenumber axis A_V for wavenumber calibration as illustrated in Fig. 6.2b. It is necessary to first record a background spectrum, B, of equivalent duration and to subtract such that $S' = S - B$. The corrected spectrum is given by $[A_V, S' \times C]$.



(a) Proposed new method of intensity and wavenumber calibration in Raman spectroscopy. In this case wavenumber calibratin is a requisite first step. Following this, spectra from two samples are required for intensity calibration: an *arbitrary uncalibrated* white light source and a reference Raman spectrum such as from glycerol.



(b) There are four steps that make up the proposed method of intensity calibrating. The requisite first step is wavenumber calibration. See text for further details.

Figure 6.3: New calibration flowchart

6.4 Intensity Calibration using an Arbitrary White-Light

The intensity calibration method described in the previous section can be problematic. The lamp requires frequent and expensive calibration and unexpected miscalibration may lead to erroneous results. In this section, we propose a novel protocol to perform intensity calibration in Raman Spectroscopy using an arbitrary white light, such as that from the microscope itself as illustrated in Fig. 6.3a. The ready availability of this light source also improves the overall applicability of intensity calibration. The method requires (i) the availability of a (true) reference Raman spectrum from a material with a known Raman spectrum, $[V_0, R_0]$ (where R_0 denotes the calibrated Raman spectrum of some appropriate material over a known wavenumber axis V_0) and (ii) the availability of an arbitrary inexpensive Tungsten Halogen white-light source such as that from the microscope itself.

Ideally, the spectrum R_0 would be highly reproducible and have non-zero intensity across the entire bandwidth. If such a material existed, there would be no need for a white light source since the correction factor for the system would simply be given by $C = R_0'/R'$, where R' denotes the raw spectrum of the material from the Rama spectrometer to be calibrated (following background subtraction), and R_0' is the value of R_0 interpolated over the same wavenumber axis as R'. However, the presence of low or negligible intensity values in regions of most Raman spectra makes this approach difficult. Some authors have proposed this very approach using materials with sharp well defined highly reproducible spectral lines in the Raman spectrum such as quinine. [59]. The premise of the approach is to avoid the bands of zero intensity in the calculation of R_0'/R', and therefore, C is calculated only for a small number of discrete points in the wavenumber axis from which the values at other wavenumber positions can be interpolated. However, this approach assumes that C will be slowly varying across the full bandwidth; this assumption is not valid for many diffraction gratings, which can exhibit highly non linear diffraction efficiency as a function of wavelength. Another difficulty with this approach is that the relative intensities of sharp spectral lines from a given material can vary significantly across spectra recorded from spectrometers with

varying resolutions, owing to the convolution of the true spectral intensity with the point-spread function of the spectrometer.

In this chapter, we investigate glycerol and a commercial plastic as reference materials to provide R_0. Both of these samples do not contain significantly sharp spectral lines, which can present problems as described above. To overcome issues relating to areas of low intensity in the spectrum, the proposed method also employs a white light spectrum from the microscope, W, which is known to have non-zero values over the full bandwidth. Unlike for the case of traditional white-light calibration, the true spectrum of this source is not known; however, this can be estimated using the values of R_0 as described in the procedure below, which is illustrated in Fig. 6.6 and Fig.6.7.

1. The first step in the proposed protocol is to perform wavenumber calibration as described in Section 6.3.2. This requires the use of a reference wavenumber standard such as 4-acetamidophenol or benzonitrile. Unlike classical intensity calibration using a calibrated white lamp, which was entirely independent of wavenumber calibration, our method requires that wavenumber calibration be implemented first. The reader may wonder if this spectrum, V, could perform the role of the aforementioned reference Raman spectrum, R, which would reduce the number of steps that follow. We have found that significantly better intensity calibration results are produced using a smoother reference spectrum such as that from glycerol.
2. In order to generate a reference Raman spectrum for the purpose of intensity calibration, it is necessary to record the spectrum, R, from the reference material as well as the background spectrum, B; these are subtracted to provide $R' = R - B$. Given the preceding step this spectrum is wavenumber calibrated and can be written as $[A_V, R']$.
3. The third step is to record the spectrum from any arbitrary white-light source such as from the microscope lamp. The true spectral irradiance T from an incandescent lamp has been discussed in Section 6.2. Over the wavelength range of most Raman spectrometers the shape of T is approximately linear or is described accurately by

a polynomial of second order. The spectrum of the microscope lamp is recorded and is denoted by W; a matching background,B, is recorded and subtracted to produce the spectrum W'. Note that the wavelength axis that corresponds to the spectrum W' is not required.

4. The fourth step is to calculate the correction factor C using the two recorded spectra W' and $[V, R']$, as well as the true reference spectrum, $[A_0, R_0]$ which is first interpolated to provide, $[A_V, R'_0]$, where R'_0 denotes the intensity values of the reference over the same wavenumber range as R'. The details of the algorithm that we propose to do this are provided in the following section. Once C has been obtained, any Raman spectrum recorded from the system can be calibrated as described in Section 6.3.3 and Fig. 6.2b.

6.4.1 Algorithm to calculate the correction factor

The goal of this algorithm is to find the correction factor, C, for which $R'_0 \approx C \times R'$. As mentioned earlier, it is not possible to simply calculate $C = R'_0 / R'$ as several problems exist with this approach listed below:

- R' will typically have large bands of zero intensity which makes division difficult.
- R'_0 and R' can have different baselines (resulting from auto-fluorescence, Mie scattering or some other form of photoluminesence), which are typically described by variable low order polynomials. [143] The relationship between R' and R'_0 is therefore more generally described as:

$$R'_0 + P_1 = R' \times C \tag{6.3}$$

  where P_1 is a low order polynomial that accounts for the difference in baseline between R' and R'_0
- If the two spectra R'_0 and R' have different resolution, this may significantly affect the spectral line shapes and may render the equality in Equation 6.3 to be inaccurate. This problem can be mitigated by using a reference sample with spectrum R_0 that does not contain sharp spectral lines; smoother spectra will be more robust

to the blurring associated with lower resolution systems. A more accurate method to deal with this problem is proposed in the discussion section.

The baseline in Raman spectroscopy can be variable due to several reasons, including:

- Fluorescence: Fluorescence emission from the sample can contribute to the background signal and cause a variable baseline. Fluorescence often arises from impurities or other materials in the sample, and can be reduced by using appropriate excitation wavelengths or by filtering out the fluorescence signal.
- Sample inhomogeneity: Samples can have areas with different optical properties, resulting in varying intensities of the Raman signal and a variable baseline. This can be addressed by ensuring that the laser spot size covers a representative area of the sample and by averaging multiple measurements.
- Instrument noise: Instrument noise, such as electronic noise, can also cause a variable baseline. This can be minimized by optimizing the measurement conditions, such as reducing the laser power or optimizing the detector settings.
- Scattering from the substrate: The substrate on which the sample is placed can also contribute to the background signal and cause a variable baseline. This can be addressed by using substrates with low Raman scattering or by using a substrate that does not interfere with the Raman spectra.
- Drift: Temperature changes or mechanical instability in the measurement setup can cause drift in the baseline over time. This can be minimized by ensuring that the measurement conditions are stable and by measuring the baseline periodically to correct for any drift.

Overall, it is important to carefully consider the measurement conditions and potential sources of variability when interpreting Raman spectra and to take appropriate measures to minimize the impact of these factors on the baseline.

The solution to the difficulty in calculating R_0'/R' is to include the white-light spectrum from the microscope lamp. The true spectrum of the white light is not known but can be approximated by a low order polynomial. Therefore, the recorded white light

spectrum from the system, W' can be described in terms of C and P_2 as follows:

$$C = P_2 / W' \tag{6.4}$$

Substituting Equation 6.3 into Equation 6.4 provides a simple expression:

$$R'_0 = P_2 \times R' / W' - P_1 \tag{6.5}$$

For low orders of P_1 and P_2 given by N_1 and N_2, respectively, it is possible to use the method of ordinary least squares to determine values for the $N_1 + 1$ coefficients of P_1 and the $N_2 + 1$ coefficients of P_2. The Matlab code for this is provided in the appendix, where the orders N_1 and N_2 can be selected as input arguments. For the baseline we have found a third order polynomial is sufficient and we set $N_1 = 3$ for all of the results presented in this chapter. To model the white light spectrum we investigate both a first order ($N_2 = 1$) and second order ($N_2 = 2$) polynomial for P_2 and have found that both cases provide high quality results.

6.5 Methods

6.5.1 Recording of Spectra

An Horiba Jobin Yvon LabRAM HR 800 Spectroscopy system was used to record spectra from a number of samples for subsequent calibration and testing. In total, two excitation laser wavelengths were used and two gratings, which amounts to four different systems. The two lasers were a 785 nm CLDS point mode diode laser with power of 300mW and a 532 nm solid state diode laser with a power of 50 mW and the two gratings were 300 lines/mm and 600 lines/mm. Given the significantly larger bandwidth and resolution of the (532 nm - 300 lines/mm) system when compared with the other three, all of which lie within the fingerprint region, this system was omitted from the analysis. This facilitated straightforward comparison of calibration results from the three systems, which had similar bandwidth and resolution. The various parameters of the three systems used to record spectra are summarised in Table 6.2. The resolution of equivalent systems with 532 nm and 785 nm excitation differs significantly [144]; in order to further simplify the comparison of calibration results across the three systems, the confocal aperture was varied such that the effective resolution of all three systems was approximately the same. The confocal aperture size and resultant resolutions are provided in Table 6.2.

In total two acquisitions were recorded from each sample in order to facilitate cosmic ray removal using the algorithm defined in Ref. 145. In summary this method compares two subsequently captured spectra and averages them except at points where the intensity varies by more than the noise floor; in such case the lower intensity value is taken. The acquisition time was chosen to be just less than the time of saturation in order to maximise the signal-to-noise ratio in each spectrum. The acquisition time that was used varied across material depending on this saturation point. In all cases a background spectrum of equivalent duration was recorded and subtracted from the sample spectrum.

Laser	532 nm	785 nm	
Laser Power	50 mW	300 mW	
Grating	600 l/mm	300 l/mm	600 l/mm
Accumulations	2	2	2
Microscope Objective	x10	x10	x10
Confocal aperture	100	200	100
Effective resolution (cm^{-1})	7.7	6.5	6.8

Table 6.2: Parameters used in spectral recording.

6.5.2 Materials and Lamps

In total, spectra were recorded from eight different chemicals on the three systems described in the previous section. These chemicals were in powder form: glucose, 4-acetamidophenol, urea and lactic acid; liquid form: benzonitrile, ethanol, and glycerol; and a solid form: polymer slide from (µ-Slide I Luer; Ibidi GmbH). We note that this polymer material has previously been investigated as a reference sample for Raman wavenumber calibration in Chapter 3. All chemicals, except for the polymer slide, were purchased from Sigma-Aldrich, Ireland.

For wavelength calibration, a Neon lamp (60910; Edmund Optics, GmbH) and a Krypton lamp (60915; Edmund Optics, GmbH) were both used to calibrate each system as described in Chapter 3. For wavenumber calibration, 14 peaks from acetaminophen were used to calibrate each system as described in Chapter 4. For traditional white light calibration, a NIST calibrated lamp was used (LS-1-CAL-INT; Ocean Optics), which was integrated with an integrating sphere (FOIS-1; Ocean Optics), which was coupled into a broadband optical fiber (QP600-2-VIS/BX; Ocean Optics). The output spectrum of this fiber was calibrated (Ocean Optics) using a NIST-traceable light source; the spectrum is sampled at a number of wavelength values and can be interpolated to provide the full irradiance over the wavelength axis of interest. For the proposed calibration method, the microscope lamp from the Bx41 microscope that is a component in the spectroscopy system was used; all filters in the illumination path were removed. In all cases a background spectrum of equivalent duration was recorded and subtracted from

the white-light spectrum.

6.5.3 Evaluation

In order to evaluate and compare the performance of the different methods for intensity calibration, we devised a quantitative metric. This metric calculates the mean absolute error (MAE) between the intensity calibrated spectrum recorded from a given material on the (532 nm - 600 lines/mm) system and the intensity calibrated spectra recorded from the other two systems with 785 nm excitation. Thus for each material, we calculate two numbers that gauge the performance of each intensity calibration method.

The mean absolute error between two intensity calibrated spectra is defined as follows:

$$MAE(X_{532}, X_{785}) = \frac{1}{M}\sum_{i=1}^{M} |X_{532}(\upsilon_i) - X_{785}(\upsilon_i)| \tag{6.6}$$

Here, X_{532} and X_{785} are both 1D vectors that contain the values of the two intensity calibrated spectra at the discrete wavenumber values given by υ_i. In total there are M samples in these two spectra. It is essential that the two spectra are defined at the same wavenumber values. It is, therefore, necessary as a first step to interpolate X_{785} onto the same wavenumber axis as X_{532}. The number of samples that overlap will determine the value of M. Two difficulties that present when using this metric are (i) the presence of differing baselines in the two spectra caused by fluorescent signals or stray light due to Mie scattering, [144] and (ii) the different relative scaling of both spectra. Both problems are solved by using a variant of the the Extended Multiplicative Scattering Correction (EMSC) Algorithm defined in Ref. 143. The goal of the algorithm is to subtract a low order polynomial baseline from the X_{785} spectrum such that the baselines for both spectra are closely matched, and secondly to scale the X_{785} spectrum such that the amplitudes of both spectra are closely matched. The basis of the approach is a least squares algorithm that can be understood as a minimisation of the error function defined in Equation 6.7 as follows:

$$error(\upsilon_i) = X_{785}(\upsilon_i) - \left(A \times X_{532}(\upsilon_i) + \sum_{m=0}^{N} a_m \upsilon_i^m \right) \tag{6.7}$$

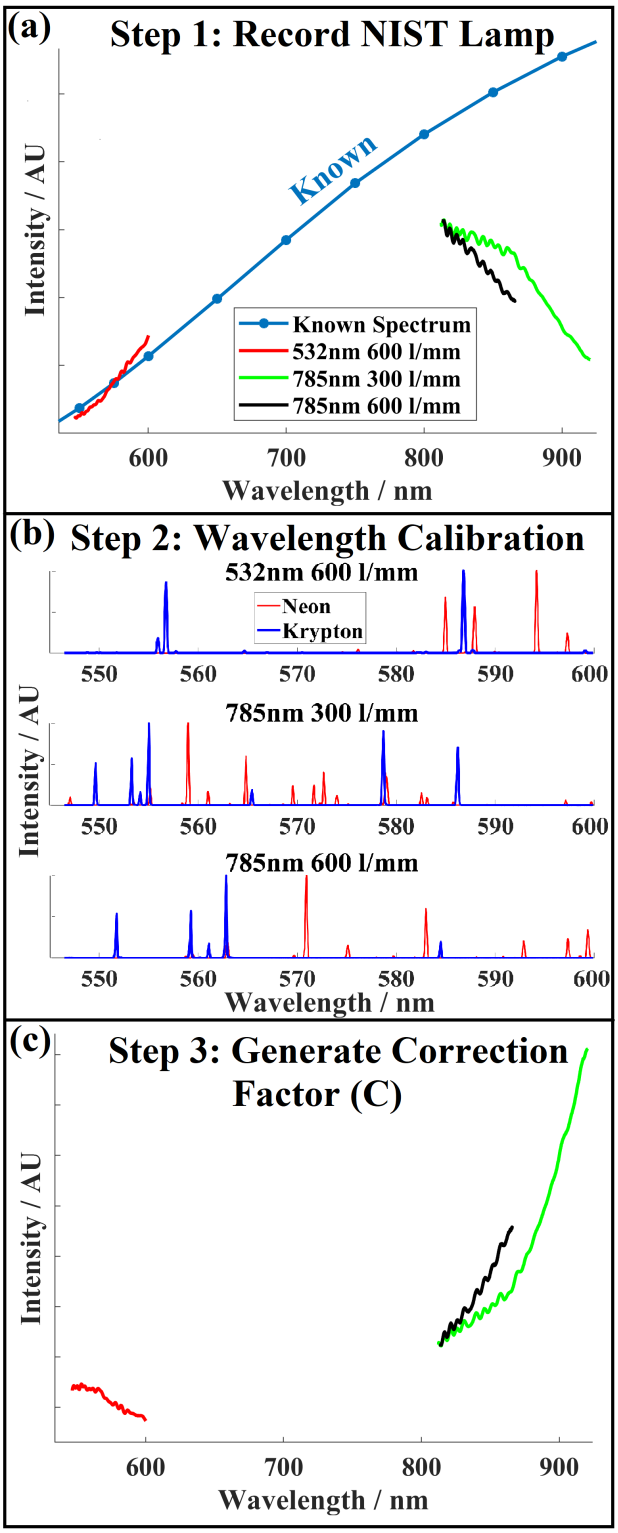


Figure 6.4: Results of the three steps for traditional intensity calibration using a known white-light source: (a) raw spectra recorded from the lamp overlaid with the known spectrum; (b) neon and krypton spectra atomic emission spectra(the properties of these two noble gas are similar, and they are all colorless and odorless monatomic gases at normal temperature and pressure, and it is difficult to carry out chemical reactions.) spectra recorded for wavelength calibration and (c) the resulting correction factors.

The goal of the algorithm is to find the weights A and $a_m : m = 0 \rightarrow N$, which minimise the square of this error function. In summary, the algorithm takes in as function arguments, X_{785}, X_{532}, and the integer N, and returns the weights described above. In our analysis, a value of $N = 2$ was used to avoid any possibility of overfitting that might be caused by the use of higher order polynomials. The Matlab code for this algorithm is provided in Appendix 6.9. Once the weights have been determined, the spectrum X_{785} can be baseline subtracted and normalised (with respect to X_{532}) as follows:

$$X'_{785} = \frac{X_{785} - \sum_{m=0}^{N} a_m v_i^m}{A} \tag{6.8}$$

The performance metric, $MAE(X_{532}, X'_{785})$, can now be calculated and used as an estimate of the quality of the intensity calibration method that produced the two spectra X_{532} and X_{785}. For each material two such numbers are calculated, one for each of the two 785 nm excitation systems. One final point is that the spectrum X_{532} should be first normalised before application of Equations 6.7, 6.8, and 6.6 such that the MAE results can be quantitatively compared across the different intensity calibration methods.

6.6 Results

6.6.1 Correction factors using known white-light

The results of the three steps (see Fig. 6.2b) involved in traditional white-light calibration are shown in Fig. 6.4. The raw white-light spectra (with background subtraction) recorded for the calibrated white light source are shown for the three systems in Fig. 6.4a. Also shown in this figure is the known spectrum of the calibrated white-light source. The distortion of this spectrum by the sensitivity response of the systems is clearly evident, particularly for the two 785 nm excitation cases. In Fig. 6.4b the Neon and Krypton spectra are shown for each case. There were sufficient peaks available from both lamps combined in all three cases to obtain a wavelength calibration accuracy of < 0.02 nm over the bandwidth of all three cases using the method outlined in Chapter 3. The final correction factors for the three systems are shown in Fig. 6.4c.

6.6.2 Correction factor using arbitrary white-light

The results of the four steps (see Fig. 6.2b) involved in the proposed arbitrary white-light calibration are shown in Fig. 6.5. The raw spectra recorded from the 4-acetamidophenol wavenumber reference material are shown in Fig. 6.5a. The peak positions were used to implement wavenumber calibration as described in Chapter 5 providing an accuracy of < 0.3 cm over the bandwidth for all three cases. The raw Raman spectra of glycerol (with background subtraction) are shown for the three systems in Fig. 6.5b. The raw white-light spectra (with background subtraction) recorded from the uncalibrated microscope lamp is shown for the three systems in Fig. 6.4c. Also shown in this figure is the known spectrum of the calibrated white-light source for comparison. In this case, however, we do not have any knowledge of the true spectrum of the lamp. The final correction factors for the three systems are shown in Fig. 6.4d having been calculated using the lamp spectrum and glycerol spectrum as described in Section 6.4.1. We note the strong similarity in the form of these correction factors with respect to the same results from the calibrated white light shown in Fig. 6.4c. Although similar there are some subtle differences, which results in more accuracy when using the microscope

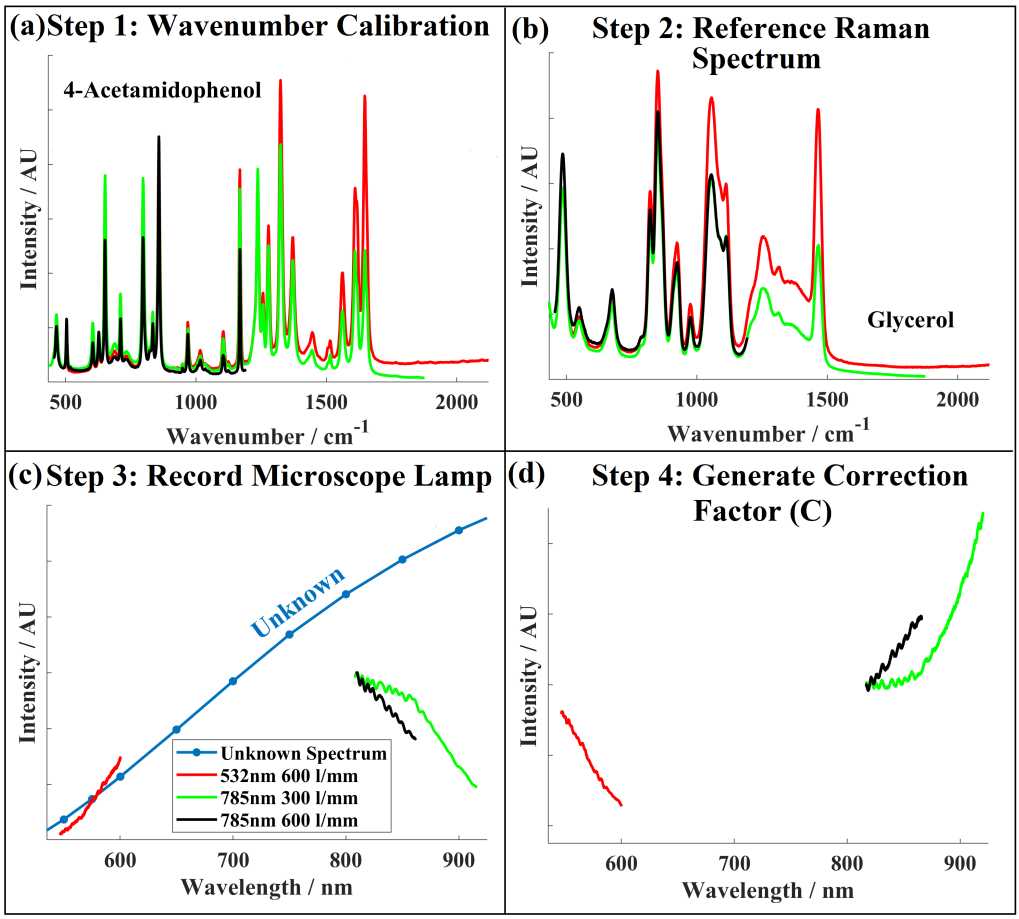


Figure 6.5: Results of the four steps for the intensity calibration method proposed in this chapter: (a) the spectra of 4-acetamidopenol used for wavenumber calibration of the three systems; (b) wavenumber calibrated raw glycerol spectra; (c) the spectrum of the unknown microscope lamp recorded using the three systems, overlaid with the same known spectrum shown in Fig. 6.5(a) for comparison; (d) the correction factors generated using the proposed method for $N_1 = 3, N_2 = 1$, see Section 6.4.1 for further details.

lamp for most cases.

6.6.3 Comparison of calibration methods

In this section, the results of three different intensity calibration methods are presented across eight different chemical spectra. These methods are (i) traditional intensity calibration using a calibrated white light source, (ii) calibration using the method proposed in this chapter with the microscope lamp and a reference Raman spectrum from glycerol, and (iii) calibration using the method proposed in this chapter with the microscope lamp and a reference Raman spectrum from a commercial polymer. The results presented here are for the case of approximating the white light profile (P_2) as

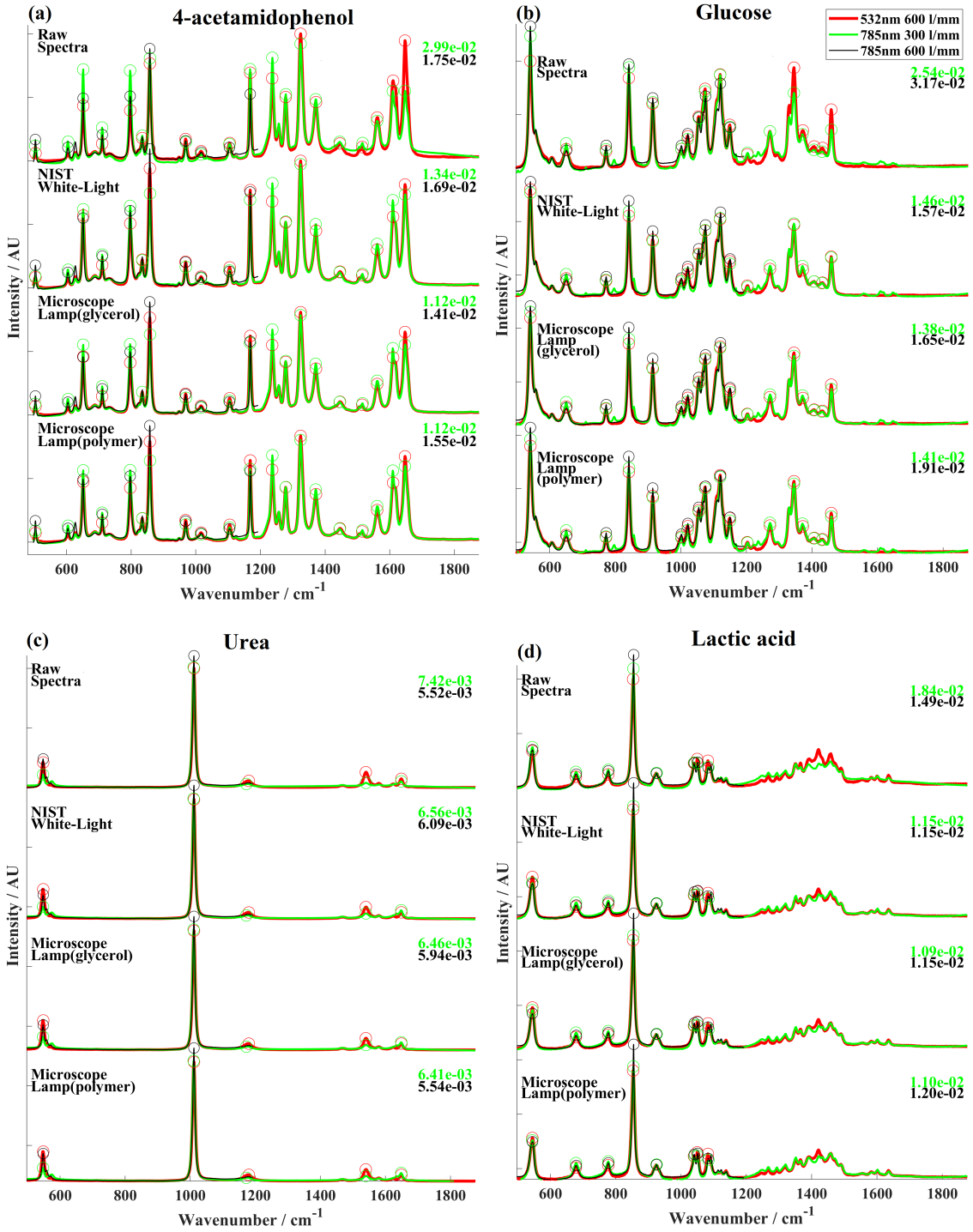


Figure 6.6: Results of the various intensity calibration protocols applied to four powdered chemicals (a) 4-acetamindophenol; (b) glucose; (c) urea; and (d) lactic acid. In all cases the red spectrum is the raw or calibrated spectrum for 532 nm excitation. For all cases the 785 nm excitation spectra have been fitted to the corresponding 532 nm spectrum in terms of normalisation and baseline subtraction, in order to facilitate quantitative comparison using the MAE as described in Section 6.7. MAE results are shown at the right of each spectrum.

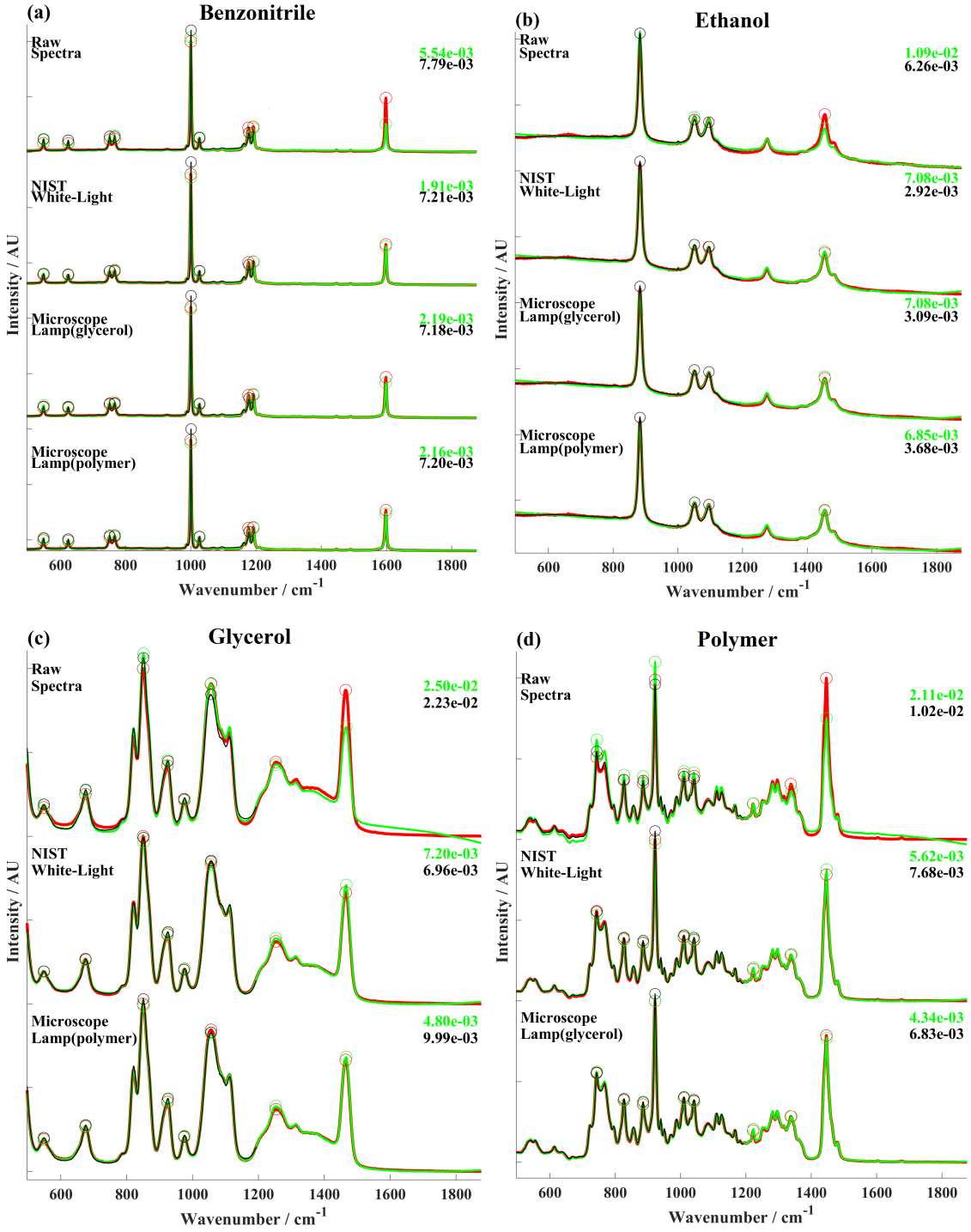


Figure 6.7: Results of the various intensity calibration protocols applied to four liquid/solid chemicals (a) benzonitrile; (b) ethanol; (c) glycerol; and (d) polymer. In all cases the red spectrum is the raw or calibrated spectrum for 532 nm excitation. For all cases the 785 nm excitation spectra have been fitted to the corresponding 532 nm spectrum in terms of normalisation and baseline subtraction, in order to facilitate quantitative comparison using the MAE as described in Section 6.7. MAE results are shown at the right of each spectrum.

a straight line, i.e. $N_2 = 1$ and the baseline (P_1) as a third order polynomial, i.e $N_1 = 3$. The raw and calibrated spectra for the three different systems (532nm-600lines/mm), (785nm-300lines/mm), and (785nm-600lines/mm), are shown in Fig. 6.6 for four solid powder chemicals: (a) 4-acetamidophenol, (b) glucose, (c) urea, and (d) lactic acid. For all cases the raw spectra recorded by the three systems show obvious differences in intensity due to the differing sensitivity responses. In order to highlight these differences, circular markers have been added to the peak maxima.

Following calibration with each of the three methods, there is clearly closer agreement between the intensity values of the spectra. Qualitatively, there appears to be little difference between the performance of the three calibration methods for all four chemicals. However, quantitatively, the MAE evaluation metric that was defined in section 6.5.3 indicates that the proposed intensity calibration method with glycerol outperforms traditional white-light calibration in all but one case: (glucose, 785nm excitation - 600 lines/mm), and in that case there were clear issues with the quality of one of the recorded spectra with unexpected peaks that we believe can be attributed to a background light source(the rest experiments data have contained enough information to support the experiment keep going). The proposed intensity calibration method with the commercial polymer also outperforms traditional white-light calibration in all but two cases: (glucose, 785nm excitation - 600 lines/mm) and (lactic acid, 785nm excitation - 600 lines/mm).

The raw and calibrated spectra for the three different systems are shown in Fig. 6.7 for three liquid chemicals: (a) benzonitrile, (b) ethanol, (c) glycerol and (d) for the solid polymer. As for the powder chemicals, the raw spectra recorded by the three systems have clear differences in intensity and following intensity calibration, these differences are appreciably reduced. Once again, the results of calibration using the three methods produce qualitatively similar results. Quantitatively, calibration with the microscope lamp and glycerol produce very similar MAE values when compared with known white light calibration; the latter performs slightly better for the case of (benzonitrile, 785nm excitation - 300 lines/mm) and (ethanol, 785nm excitation - 600 lines/mm), while the

former is slightly more accurate for the other cases, with the exception of the final two cases (polymer, 785nm excitation - 600 lines/mm) and (polymer, 785nm excitation - 300 lines/mm) for which the microscope lamp/glycerol method is significantly more accurate. Calibration with the microscope lamp and polymer produces mixed results. On the one hand the MAE results are similar or better than known white light calibration for some cases eg. (ethanol, 785nm excitation - 300 lines/mm) and (glycerol, 785nm excitation - 300 lines/mm) and for other cases the results are markedly worse eg. (ethanol, 785nm excitation - 600 lines/mm) and (glycerol, 785nm excitation - 600 lines/mm)

In order to gain further insight into the the relative performance of the three calibration methods, the results were tabulated (see Table 6.3 Appending) and a bar chart is provided in Fig. 6.8, which is discussed further in the next section.

6.7 Discussion

The bar chart in Fig. 6.8 reveals several interesting points. Firstly, and most importantly, the proposed method using glycerol (with $N_2 = 1$ or $N_2 = 2$) provides equivalent or better intensity calibration when compared with traditional known white-light calibration in almost all cases. We believe that the reason for this is that the known white light is not perfect; it cannot be expected to produce an exactly reproducible spectrum in all cases. This is evidenced by the fact that such lamps require frequent and costly re-calibration. As discussed in section 6.2, small changes in the temperature of the lamp, brought about by unexpected variation in the current, can result in appreciable changes in the spectrum, such that it will differ from the 'known' spectrum that is provided by the manufacturer or calibration company. The method proposed in this chapter is robust to this problem as there is no assumption made about the true spectrum of the lamp other than it can be modeled as a low order polynomial over the bandwidth of the spectrometer. A second point of note is that not all reference materials can be expected to perform well. Using the polymer as a reference material provided less consistent results when compared with glycerol. While the polymer produces equivalent or better results than traditional known white light calibration for one (785nm excitation 300 lines/mm) system the results are slightly worse for the second system. We believe that this is due to the nature of the polymer spectrum containing several sharp peaks, which can vary considerably as a function of the system resolution. This affects the utility of this spectrum as a reproducible Raman reference spectrum to be included in the algorithm presented in Section 6.4.1. As further evidence of this issue, 4-acetamidophenol, whose spectrum contains even sharper peaks than the polymer spectrum, was also tested as a reference Raman material but the results, not shown here, were poor. Glycerol, in contrast, has a spectrum that is relatively slowly varying and the shape of the spectrum is more robust to variation in system resolution. If the method could be enhanced such that the polymer or 4-acetamidophenol produced reliable results for intensity calibration, this would be advantageous as these could also be used as wavenumber reference standards in step 1 of the calibration process shown

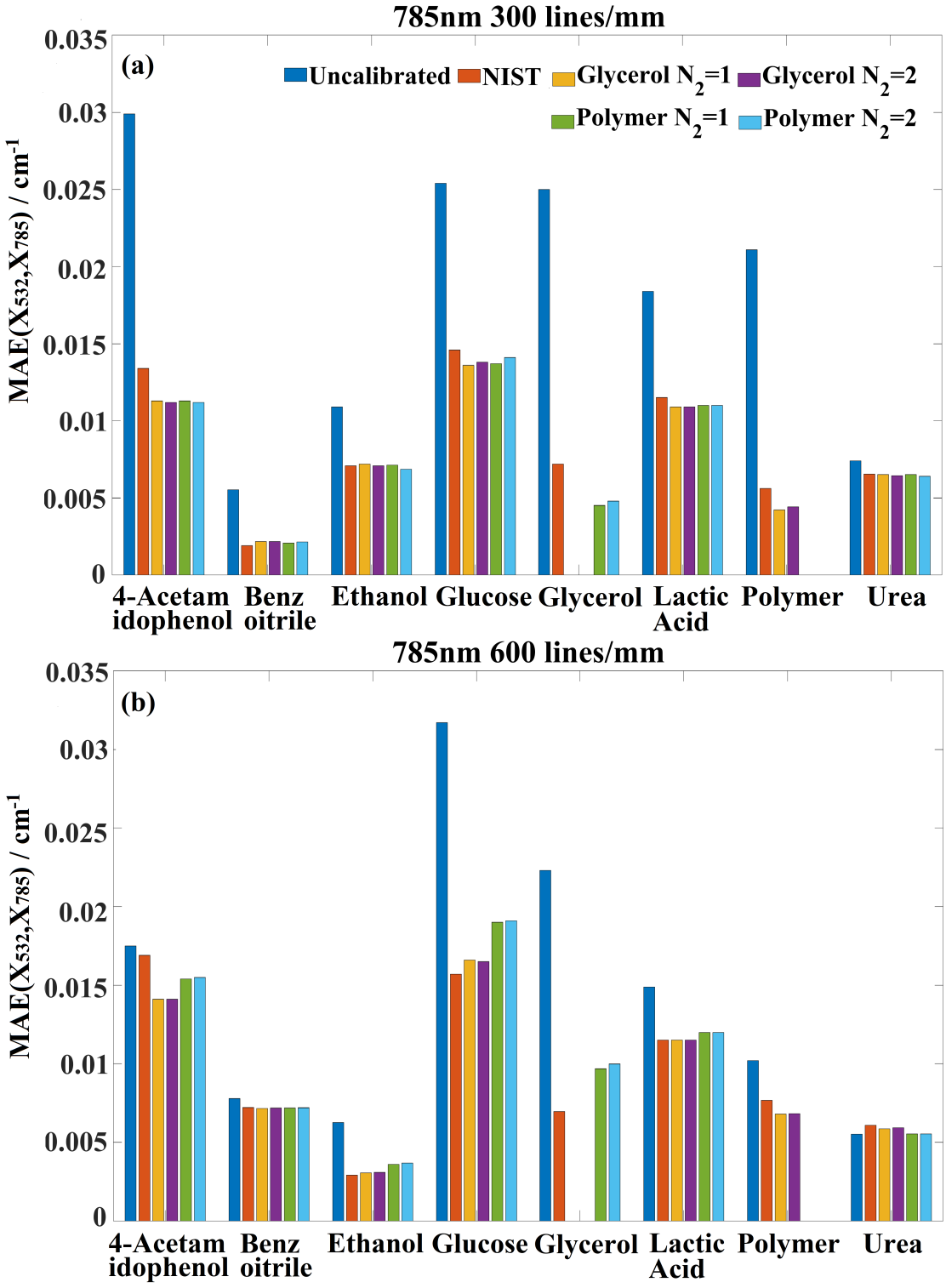


Figure 6.8: This bar chart reveals the proposed method using glycerol ($N_2 = 1$ or $N_2 = 2$) provides an equivalent or better intensity calibration than traditional known white light calibration in almost all cases.

in Fig. 6.3b, as described in Chapter 3, which would reduce the overall number of steps to three.

Another point of discussion is the disparity between the peaks in the calibrated spectra in Fig 6.6 and Fig 6.7. One may enquire why the peaks in these spectra do not perfectly match up following calibration. We believe the reasons for this are two fold: firstly differences in resolution can affect relative peak amplitudes, and these differences are more pronounced for narrow lines. Secondly, polarisation can be a major factor in the accuracy of intensity calibration. The depolarisation ratios of the various vibrational resonances that make up a single spectrum can vary significantly. In practical terms, this means that each spectral line in the spectrum can have different ratios of horizontal and vertical polarisation. Since the sensitivity response of the Raman system is polarisation dependent, owing to the different diffraction efficiency for s- and p- polarised light, varying states of polarisation across the spectrum presents a significant problem for intensity calibration methods. Solutions to this problem include: (i) performing the overall calibration routine separately for vertically and horizontally polarised light making use of a linear polariser at the input to the spectrograph. This would either reduce the intensity of recorded spectra by half or necessitate two captures for each spectrum and two calibration correction factors; (ii) Insertion of a quarter waveplate or a random polariser, which could serve to approximately equalise the light power in both orthogonal polarisations such that a single intensity calibration routine can be more accurately applied. Alternatively a mirror system could be inserted to flip the polarisation by 45 degrees to produce the same effect. [146]

One avenue for improving the proposed intensity calibration method could be the use of a Gaussian convolution applied to the reference Raman spectrum in order to match the resolution of the system undergoing calibration with that of the reference spectrum. Such an approach might render the polymer, or even 4-acetamidophenol, as a reproducible reference Raman spectra with the aforementioned advantages. This could also improve the accuracy of the method for any given reference material including glycerol. Another advantage of this approach is that there would be no requirement

to physically vary the slit/pinhole in the spectrometer such that the resolution of the reference spectrum and system undergoing calibration are matched. A further extension of the method might be to use several reference materials and to take an average of the correction factors produced using each one.

6.8 Conclusion

In this chapter, a novel method of intensity calibration of Raman spectrometers is proposed. This method is based on using an arbitrary white light with unknown spectral profile together with a reference Raman material with known Raman spectrum. The white-light spectrum combined with the Raman spectrum of the reference enables the true spectrum of the white light to be estimated, and therefore, the sensitivity response of the system to be obtained for the purpose of calibration. Inherent in this algorithm is the principle that the white light spectrum can be modelled by a low order polynomial over the bandwidth of the spectrometer. We have demonstrated that this assumption is valid based on modelling the spectral output from a Tungsten-Halogen lamp.

In order to prove the accuracy of the approach, a metric was developed that enables direct quantitative comparison between two intensity calibrated Raman spectra of the same material recorded from different systems with approximately similar resolution, which can take into account the variable baseline in both spectra. Applying this metric to spectra recorded from eight materials with 785 nm and 532 nm excitation systems, demonstrates that the proposed method, making use of a microscope lamp and glycerol reference spectrum, produces intensity calibration that is equivalent or more accurate than traditional known white-light calibration in almost all cases.

Although the results presented in this chapter are promising, further work is required to fully develop the method into an established Raman intensity calibration protocol. Further testing of the method is required over more systems and over wider bandwidths; in this chapter only the fingerprint region was analysed. Several extensions and improvements are suggested in the previous sections including development of the numerical algorithm to work with systems of any spectral resolution, which is a necessary step for ubiquitous application of the method. Given the inexpensive and simple approach of recording an arbitrary white light spectrum and a glycerol Raman spectrum, we believe the proposed protocol may offer a much needed calibration protocol to replace the current standards of known-white light calibration and fluorescence standards. The classical methods are often not applied at all due to their expense, difficulty in mounting

the sources, issues with long-term maintenance, and/or lack of transferability across systems. A simple inexpensive approach is far more likely to be applied regularly by Raman users.

The core topic of this chapter is the development of a new intensity calibration protocol. There exists a secondary contribution, which should not be overlooked; in the course of this work we proposed a novel metric that can be used to quantify the accuracy of an intensity calibration protocol across two or more instruments. Previous attempts to gauge the performance of an intensity calibration protocol have been qualitative in nature. One reason for this is the variable baseline that can occur when recording spectra from the same material on different instruments, especially with different excitation wavelengths. The metric proposed here can subtract this baseline difference and measures accuracy only on the spontaneous Raman resonance.

In this chapter, we have investigated a novel method for intensity calibration of a Raman spectrometer. In the next chapter, we return to the subject of wavelength calibration, this time making use of machine learning methods.

6.9 Appendix

Material	System (785 nm)	Uncali-brated	NIST Calibrated	Glycerol 1st	Glycerol 2rd	Polymer 1st	Polymer 2rd
ACE	785/300	2.99E-02	1.34e-02	1.13E-02	1.12E-02	1.13E-02	1.12E-02
	785/600	1.75E-02	1.69E-02	1.41E-02	1.41E-02	1.54E-02	1.55E-02
Ben	785/300	5.54E-03	1.91E-03	2.19E-03	2.19E-03	2.09E-03	2.16E-03
	785/600	7.79E-03	7.21E-03	7.15E-03	7.18E-03	7.18E-03	7.20E-03
Ethanol	785/300	1.09E-02	7.08E-03	7.20E-03	7.08E-03	7.13E-03	6.85E-03
	785/600	6.26E-03	2.92E-03	3.06E-03	3.09E-03	3.60E-03	3.68E-03
Glucose	785/300	2.54E-02	1.46E-02	1.36E-02	1.38E-02	1.37E-02	1.41E-02
	785/600	3.17E-02	1.57E-02	1.66E-02	1.65E-02	1.90E-02	1.91E-02
Glycerol	785/300	2.50E-02	7.20E-03	X	X	4.52E-03	4.80E-03
	785/600	2.23E-02	6.96E-03	X	X	9.68E-03	9.99E-03
Lactic acid	785/300	1.84E-02	1.15E-02	1.09E-02	1.09E-02	1.10E-02	1.10E-02
	785/600	1.49E-02	1.15E-02	1.15E-02	1.15E-02	1.20E-02	1.20E-02
Polymer	785/300	2.11E-02	5.62E-03	4.23E-03	4.43E-03	X	X
	785/600	1.02E-02	7.68E-03	6.80E-03	6.83E-03	X	X
Urea	785/300	7.42E-03	6.56E-03	6.52E-03	6.46E-03	6.53E-03	6.41E-03
	785/600	5.52E-03	6.09E-03	5.87E-03	5.94E-03	5.53E-03	5.54E-03

Table 6.3: Table of all results, red means the best value

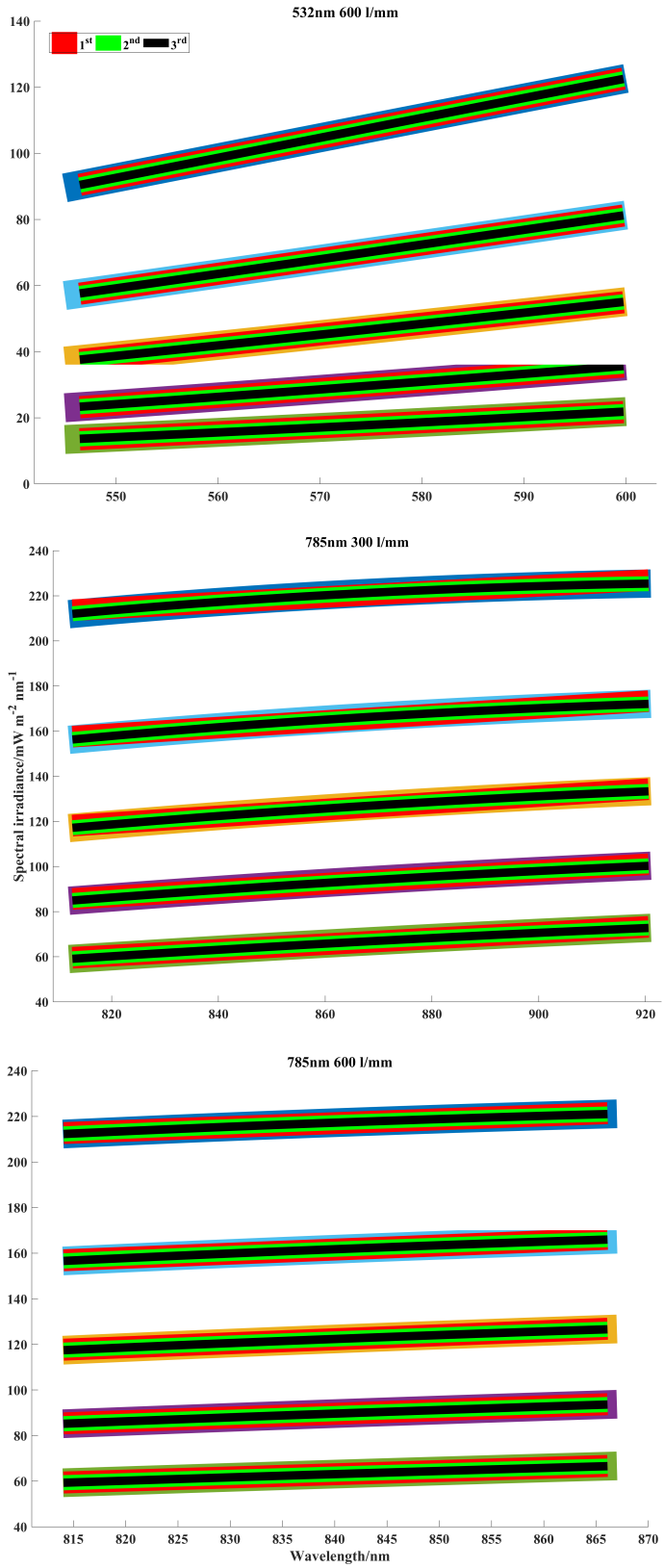


Figure 6.9: First (red), second (green), and third (black) order polynomial fits applied to the five temperature dependent lamp irradiance profiles shown in Fig. 6.1 in the bands of the three Raman spectrometers used in this study. The mean absolute error for these fits are provided in Table 6.1. While first order fitting provides low error, second order provides a closer fit with negligible improvement for third order fitting.

```
function[c,s,base,truelamp,score]=Intensity_Calibration (S,r,lamp,N1,N2)
S is the spectrum we want to correct
r is a reference spectrum that has previously been calibrated (and possibly blured
to match the system resolution.
lamp is a recording of ANY lamp with dark current subtracted
N1 should be 3
N2 should be 1
x = 1:length(S);
mu = [mean(x); std(x)];
x = (x - mu(1))/mu(2);
x = x';
V(:, N1 + 1) = ones(length(x), 1, class(x));
for j = N1 : -1 : 1
V(:, j) = x. * V(:, j + 1);
end
V(:, N1 + N2 + 2) = ones(length(x), 1, class(x)). * S'./lamp';
for j = N1 + N2 + 1 : -1 : N1 + 2
V(:, j) = x. * V(:, j + 1);
end
[Q, R] = qr(V, 0);
p = R/(Q' * r');
c_baseline = p(1 : N1 + 1);
c_lamp = p(N1 + 2 : end);
baseline = polyval(c_baseline, x);
true_lamp = polyval(c_lamp, x);
s = (S. * true_lamp')./lamp + baseline';
score = sum(abs(s - r));
```

Figure 6.10: Code for proposed method

Chapter 7

Wavelength Calibration using Long Short Term Memory Architectures

The work in this chapter has been prepared as a conference proceeding and will be submitted to an appropriate SPIE conference at the next convenience: *Liu, Dongyue, and Bryan M. Hennelly. "Wavelength Calibration using Long Short Term Memory Architectures." to be submitted to the proceedings of the SPIE in 2023* with the following abstract:

"In this paper, the long-term short-term memory (LSTM) architecture is investigated as a tool for wavelength calibration of a spectrometer. Polynomial fitting is the most common method of wavelength calibration, whereby wavelength standards, such as neon and krypton are recorded and the position of the spectral lines on the detector together with the known wavelengths are used for fitting. The method performs poorly when only a small number of lines appear within the bandwidth recorded by the spectrometer. We demonstrate how the basic encoder-decoder LSTM architecture can be used to provide superior wavelength calibration accuracy when five or less lines are present. We believe with further development, machine learning could outperform the traditional methods in all cases"

7.1 Introduction

In this chapter, we return to the subject of wavelength calibration, which was reviewed in Section 2.4 in Chapter 2 and discussed in more detail in Chapter 4. This time we investigate machine learning applied directly to the atomic emission spectrum, as a means to obtain the calibrated wavelength axis. Like all machine learning methods, the availability of a large dataset for training is essential. The derivation in Chapter 4 relating the detector to the wavelength axis plays a key role in this chapter as a component for the simulation of experimental neon spectra recorded from a variety of spectrometers with variable physical parameters such as the grating period and the slit width.

Polynomial fitting is a traditional calibration method for wavelength calibration. It is the most popular and widely used method because of its efficiency and simplicity. Wavelength standards, such as neon and krypton are well characterised and provide excellent references for wavelength calibration using polynomial fitting [58, 147, 148]. A low-order (first, second or third order) polynomial is fit to a series of pixel positions and associated known reference wavelength coordinates [39, 59, 61–63, 69–72, 98, 149–153]. Polynomial fitting, however, has limitations in terms of accuracy, which is sensitive to the spectrograph slit width, and is especially sensitive to the number, sharpness, and distribution range of the peaks in the reference standard. It is also highly sensitive to the presence of noise, which can significantly impact the perceived peak position in the reference spectrum. Another limitation of significance relates to the bands at the ends of the spectrum. If no peaks exist at the ends of the spectrum, polynomial fitting is likely to produce significant errors in these bands. This necessitates the use of reference standards that have a large number of peaks across the entire bandwidth of the spectrum, which are generally not available for many grating angles.

With the limitations of polynomial fitting in mind, we set out to investigate alternative methods. Machine learning has been the subject of rapid advancements in recent years. Deep learning [154] is a technology that can extract data features with multiple processing layers. [148] By training on extensive training sets deep learning has already been established as a useful technology for processing spectra. [155–160]. Of particular

interest in this chapter is the long-term short-term memory (LSTM) network. An LSTM is a sub class of Recurrent Neural Network (RNN), where the network is trained by looping through sequences of data as functions of time. While convolutional neural networks, which are more commonly applied in processing tasks including denoising, employ filters within convolutional layers to transform data, LSTMs/RNNs are predictive in nature, reusing activation functions from other data points in the sequence in order to generate the next output in a series. LSTM is sometimes preferred over other deep learning architectures because it is well-suited for handling sequential data with long-term dependencies. Traditional neural networks, such as feedforward neural networks, are not very effective at modeling sequences of data because they treat each input as independent and do not account for the order in which the data points are presented.

LSTMs, on the other hand, are a type of recurrent neural network (RNN) that can process sequences of data and remember past inputs over long time periods. They use a gating mechanism to selectively update and forget information from previous time steps, allowing them to effectively handle sequences with long-term dependencies. This makes them particularly useful for tasks such as language modeling, speech recognition, and time series prediction.

In addition, LSTMs can also handle variable-length input sequences and can learn to extract relevant features from the input data automatically, reducing the need for manual feature engineering. Overall, the ability of LSTMs to handle sequential data with long-term dependencies, along with their flexibility and automatic feature extraction capabilities, make them a popular choice for a wide range of deep learning applications. Given the repeatable pattern that is a calibration spectrum, we believe the LSTM is a suitable choice of architecture. LSTM also solves the gradient disappearance and explosion problems of RNN. The gradient disappearance and explosion problems of RNN are mainly caused by the cyclic multiplication of the weight matrix of RNN. LSTM can solve the above problems well through its own complex structure. In this chapter, a network architecture based on LSTMs, commonly referred to as an encoder-decoder is

investigated as a method to perform wavelength calibration. The concept is that this network could be taken in a neon spectrum recorded by an arbitrary Czerny-Turner spectrometer, and could return an estimate of the wavelength axis.

To achieve this, a large dataset is required. Given the difficulty in recording such a large and varied dataset, it is preferable to build a simulator that can accurately emulate the behaviour of the spectrometer and can produce many spectra in a short time, together with the known 'true' wavelength axis. In section 7.3, we will present a method to generate the simulated neon spectra and wavelength axis with different parameters including different spectrometer focal length, grating period, slit width, grating angle, and detector size and noise. This method is based on the physical model developed in section 4.3 in Chapter 4, which is augmented to include the effect of the slit width. We demonstrate that this method can provide superior results when compared with third order polynomial fitting, when only a few neon peaks are captured within the bandwidth of the spectrometer.

7.2 Long Term Short Term Memory - Encoder/Decoder Model

A second type of deep learning architecture was also investigated for wavelength calibration using a similar dataset based on simulated Neon spectra, based on the Long-Term Short-Term Memory (LSTM) architecture, illustrated in Fig. 7.1 which is a sub-class of the more general Recurrent Neural Network [161] (RNN), typically applied to process sequential data. LSTMs [162–164] contain feedback connections, which differ from the principle of the feed-forward/propagate-backwards used by CNNs like VECTOR; LSTMs were developed to overcome the vanishing gradient problem often encountered with RNNs and are particularly useful for processing 1-D sequences of data by learning from recurring patterns in datasets that appear sporadically in time. The connection weights and biases in the LSTM adjust with each epoch, in a similar manner to the storage of long-term memories via synaptic strengthening in the brain; whereas activation patterns in the network change once per time-step, in a similar manner to short-term memory. The LSTM, illustrated in Fig. 7.1 is composed of cells,

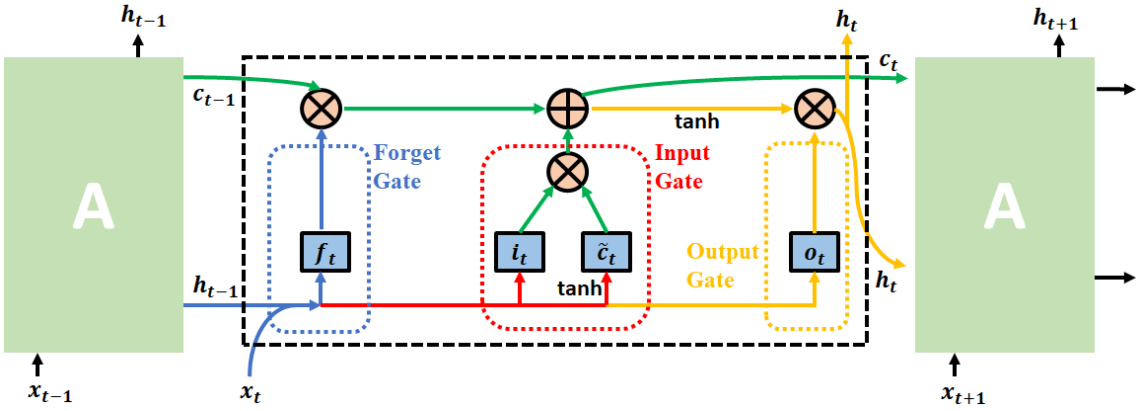


Figure 7.1: Diagram of the Long Short-Term Memory network. LSTM is formed using four main gates; the input gate, the forget gate, the output gate and the cell state. These gates are connected in a particular way to learn the long term dependencies.

which contain an input gate, an output gate and a forget gate as shown in the figure. The cell 'remembers' values over arbitrary time intervals and the three gates control the flow of information into and out of the cell. The LSTM has two transmission states, c_t

(the cell state), and h_t (the hidden state). The h_t state is passed down slowly whereby the output h_t depends on the previous state h_{t-1} plus some additional values. The c_t is more susceptible to change due to the effect of the forget gate. Using the current input into the cell x_t, and the h_{t-1} passed down from the previous state, the four layers can be trained: the forget gate f_t; the input gate i_t; a vector of new candidate values $\widetilde{c}_t$ sometimes referred to as the input modulation gate; and the output gate o_t, which are represented by blue diagrams in the figure. Therefore, f_t, i_t, o_t are often implemented using sigmoid activation functions which map the input beween 0 and 1, while $\widetilde{c}_t$ uses an *tanh* activation function to convert the result into a value between -1 and 1. The light orange circles in the figure represent matrix addition or multiplication.

There are three main stages inside the LSTM cell. The first stage is given by the blue lines and relate to the forget gate, which controls the degree to which c_{t-1} of the previous state will be 'remembered' or 'forgotten'.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{7.1}$$

where σ is the sigmoid function, $W_{f,i,o,c}$ and $b_{f,i,o,c}$ are the weight matrices and the bias vector parameters for each of the four layers, respectively. Secondly, the red lines relate to the the input gate and represent the select memory stage. This stage selectively 'memorizes' the inputs of this stage. The current input content is represented by $\widetilde{c}_t$ and the selected gating signal is controlled by i_t:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{7.2}$$

$$\widetilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{7.3}$$

The values are multiplied and added to the old cell state c_{t-1} which has been forget gated as previously described:

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \widetilde{c}_t \tag{7.4}$$

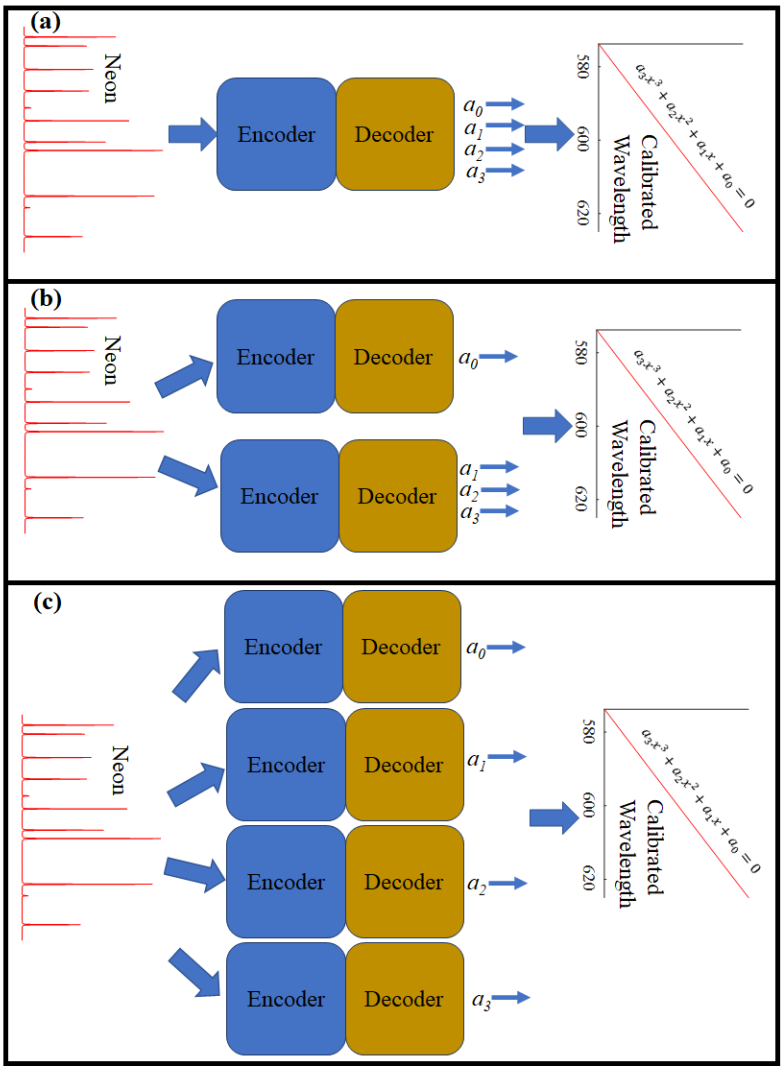


Figure 7.2: Illustration of the three networks that were investigated in this study. The first network shown in (a) contains a single encoder/decoder pair with a variable number of LSTMs in the encoder and decoder. This model was found to work poorly because the loss function prioritised the larger coefficient a_0. As a result we investigated tow other networks shown in (b) and (c). For the second network, two encoder/decoder pairs were used where the first is trained to produce the first coefficient, and the second pair is trained to produce the latter three coefficients. The third architecture uses a separate encoder/decoder pair for each coefficient.

This stage is represented as the green lines in the figure. Finally the output stage is represented by the orange lines and this stage determines the output of the current state, mainly controlled by o_t. It also scales the c_t obtained in the previous stage via a *tanh* activation function to get the h_t.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{7.5}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{7.6}$$

To solve sequence problems in which there are a different number of input and output sequence values, the encoder-decoder model has been designed, based on an architecture with two LSTM layers. It is specifically this type of architecture that we use here, where the input is the neon spectrum of length 1000 samples and the output is the four coefficients that uniquely define a third order polynomial that accurately represents the wavelength-pixel relationship. The first LSTM layer works as an encoder layer and encodes the input sequence into hidden vectors. The decoder is also an LSTM layer, which accepts three inputs: the encoded sequence from the encoder LSTM, the previous hidden state, and the current input. During training the output at each time-step is used to train the encoder-decoder model. While making predictions, the encoder output, the current hidden state, and the previous output are used as input to make the prediction at each time-step. As with the standard LSTM network, a dense layer is used to produce the output for the network.

In the first attempt illustrated in Fig. 7.2 (a) a single LSTM encoder decoder pair were trained to take a Neon spectrum as input (of length 1000) and to produce the four coefficients; the loss function used to train the network is given below:

$$\mathcal{L}_1 = \frac{1}{4} \sum_{i=0}^{3} \left| f(\mathbf{X}; \theta; i) - \mathbf{Y}_i \right|, \tag{7.7}$$

where $\mathbf{X}$ is the input neon spectrum, $\mathbf{Y}$ is the set of four coefficients that define the third order polynomial that has been fit to the true wavelength axis. This third order polynomial represents the wavelength-pixel relationship with a high level of accuracy,

$f(\mathbf{X};\theta;i)$ represents the four part tuple that is output by the network when $\mathbf{X}$ is input and i has four index values from 0 to 3 which relate to the four coefficients, θ represents all of the network parameters, which have been optimised by the training process with no a priori knowledge. However, a problem presented whereby the values of $f(\mathbf{X};\theta;i)$ for $i = 0 \rightarrow 3$ had a significantly different range of values. Going forward we will refer to these values as (a_0, a_1, a_2, a_3), which are the coefficients for $(1, x, x^2, x^3)$, respectively. For the neon spectra that were processed we found that a_0 would have a value of approximately 0.5 (units of μm), while a_1, a_2, a_3 would have smaller values of approximately $1 \times 10^{-4}, 1 \times 10^{-4}, 1 \times 9^{-13}$. This created a problem when using the MAE as defined in Equation 7.7, whereby an error in a_0 carried significantly more weight in the loss function than errors in the other four variables. While accurate results were obtained for the value of a_0, poor results were obtained for the other variables. The decision was taken to try two LSTM encoder/decoder pairs as illustrated in Fig. 7.2(b), whereby the first LSTM was designed/trained to take the Neon spectrum as input and produce a single output in the form of the largest coefficient a_0, and the second LSTM encoder/decoder pair was trained to take the Neon spectrum as input and produce a single output in the form of the other three coefficients a_1, a_2, a_3. The loss functions for these two networks are defined as follows:

$$\mathcal{L}_1 = \left| f(\mathbf{X};\theta;0) - \mathbf{Y}_0 \right|, \mathcal{L}_1 = \frac{1}{3} \sum_{i=1}^{3} \left| f(\mathbf{X};\theta;i) - \mathbf{Y}_i \right|, \tag{7.8}$$

A third design was also investigated using four LSTM encoder/decoder pairs, where each one takes the neon spectrum as input and outputs a single coefficient. The loss function for each case is similar to the first part of Equation 7.8. Results for the latter two networks are provided in Section 7.4.

7.3 Training Sets: Simulating a Neon Spectrum

In order to generate a deep learning model that can predict the wavelength or wavenumber axis from a recorded (uncalibrated) reference spectrum, it is necessary to generate a large dataset of pairs of one dimensional vectors in the form of a reference spectral intensity (such as neon) and its corresponding wavelength axis as might be expected from a real world spectrometer. In this subsection, we discuss how this is accomplished for the case of a neon spectrum and a typical Czerny-Turner Spectrograph with a variable slit aperture. We begin by reviewing the ideal neon spectrum in terms of its known emission lines as defined by the National Institute of Standards and Technology (NIST), and how these lines can be modelled using Lorenzian functions. This is followed by a discussion of how the spectrograph can be modelled in terms of the basic optical components such that the spectral intensity pattern can be predicted on the array detector, which is subject to change with variation in the parameters of the optical system such as the diffraction grating period and/or angle as well as the slit aperture width.

7.3.1 Ideal Neon Spectrum

In this section, we review the reference Neon spectrum in terms of the known emission lines. Emission occurs at particular wavelengths relating to transitions in quantum states caused by electrical stimulation. A list of well known emission lines are provided in Table 7.1 and an example spectrum of Neon recorded on an Andor Shamrock 500 spectrometer with 300 lines/mm is shown in Fig. 7.3(b). Also shown in the figure are the bands that are recorded by two other dispersive gratings, a 600 lines/mm and a 1000 lines/mm grating. An Andor Idus camera was used to record the spectrum shown in this figure. We note that the three spectral windows shown in Fig. 7.3 in the black, blue, and red boxes correspond to specific values of θ_d, the grating rotation angle, which were chosen arbitrarily when recording the spectra; f, the focal length of the spectrograph, and α, the half deviation angle are fixed for a given spectrometer.

The spectral line shape of each emission line in the reference Neon spectrum can be

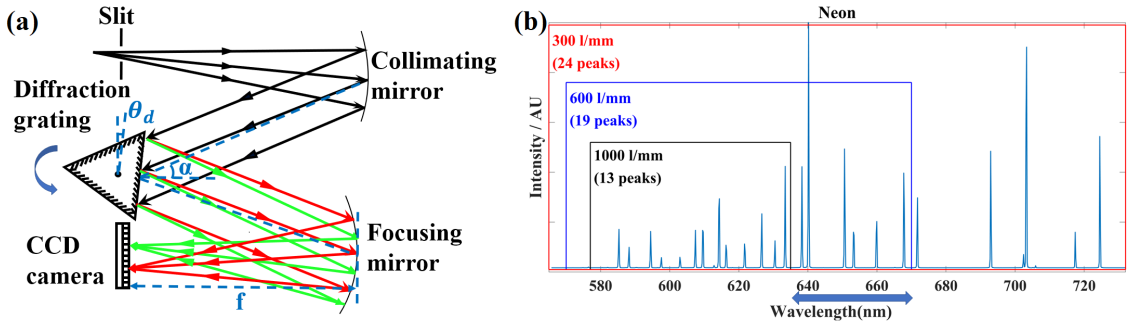


Figure 7.3: (a) The common Czerny-Tuner spectrometer, which is the basis of the calibration algorithm developed in this chapter. We note, however, the methods proposed here can be extended to other spectrometer architectures if a mathematical model is availble to relate wavelength to pixel position on the detector; (b) Experimentally recorded neon spectra recorded using a 300 lines/mm grating in a 0.5 m focal length Czerny-Turner Spectrometer. A slit width of 25 μm was used which is equal to the width of the pixel in the detector. Highlighted in the figure are the bands that would be recorded by three different dispersive gratings for arbitrary value of θ_d.

ideally modelled using a Lorentzian function [165, 166] as described by the equation below:

$$f(\lambda,\lambda_i) = \frac{(\frac{1}{2}\Gamma)^2}{(\lambda-\lambda_i)^2+(\frac{1}{2}\Gamma)^2} \tag{7.9}$$

where the centre is denoted by λ_i, and full width half maximum is given by Γ. An ideal Neon spectrum can, therefore, be modelled by a superposition of several such Lorentzian functions as defined by the equation below:

$$Neon(\lambda) = \sum_{i=1}^{24} A_i f(\lambda,\lambda_i) \tag{7.10}$$

where A_i denotes the relative intensity or amplitude of the given line. We note that for this model, we have assumed that all of the peaks have the same full width half maximum, which may not necessarily be true. Furthermore, these widths are subject to change depending on temperature, pressure and phase [165, 166]. However, we have found that modelling the spectrum to contain peaks all of line width 0.01 nm to 0.02 nm is sufficiently accurate for the purposes of our work. It appears that selecting a line width less than this provides no difference in the simulated spectra that are discussed in section 7.3.4 because the slit function is the dominating factor in producing the actual peak width in the intensity pattern at the detector; further, selecting a Lorentzian width

Peak No.	Wavelength / nm	Relative Intensity
1	585.24878	20000
2	588.18950	10000
3	594.48340	5000
4	597.55343	6000
5	602.99968	10000
6	607.43376	10000
7	609.61630	3000
8	614.30627	10000
9	616.35937	10000
10	621.72812	10000
11	626.64952	10000
12	630.47893	1000
13	633.44276	10000
14	638.29914	10000
15	640.22480	20000
16	650.65277	15000
17	653.28824	1000
18	659.89528	10000
19	667.82766	5000
20	671.70430	700
21	692.94672	100000
22	703.24128	85000
23	717.39380	77000
24	724.51665	77000

Table 7.1: List of emission lines for neon taken from Reference 1 in the range 585 nm to 725 nm. Also shown in the Table are the approximate relative intensities of the peaks as provided by NIST[1]

less than this range also forces a high level of sampling which becomes computationally intractable. The role of the slit function is discussed further in Section 7.3.3.

7.3.2 Modelling the Spectrometer

In this section, we briefly review the traditional Czerny-Turner Spectrometer, and discuss how this architecture can be modelled such that the spectral intensity incident on the array detector can be predicted, given a known light field input to the slit. This work is based on the derivation in Section 4.3 in Chapter 4. In summary, the parameters of interest that define the relationship between the wavelength axis λ, and the detector pixel axis(x, include the diffraction grating period d, the focal length of the parabolic mirrors f, the pixel width of detector T, the grating rotation angle θ_d, the angle of the

optical axis α, the misalignment of the centre of the detector array with respect to the optical axis C. The relationship between wavelength and pixel is given as follows:

$$\lambda = \frac{d}{n}\left\{\sin\left[\tan^{-1}\left(\frac{xT-C}{f}\right)+\alpha-\theta_d\right]+k\sin\left(-\alpha-\theta_d\right)\right\} \tag{7.11}$$

A typical Czerny-Turner architecture such as the one illustrated in Fig. 7.3 (a) will in general be provided from the manufacturer with a set of specifications that define the values of the aforementioned parameters; however, the accuracy of these parameters is rarely provided. Moreover, some of these parameters cannot be known with high accuracy; examples of this are the angle α defined by the angle of the optical axis, which can be estimated using a protractor or similar method of measuring angle; the angle of rotation of the grating, θ_d, which is commonly variable; and the slit width, which is also variable, and which is covered in the subsection that follows.

7.3.3 Varying The Slit Width

The point spread function (or impulse response) of the imaging system that is the spectrometer, i.e. the two parabolic mirrors, will have a point spread function that can be approximated by a Gaussian function; as such we can assume that a ideally narrow slit will produce an image on the spectrograph that is the neon spectrum convolved with this point spread function. In general this Gaussian function has a width that is smaller than the pixel size of the detector and we can ignore the effect. More interesting is the effect of slit width on the image when the slit is not ideally narrow. Assuming that there exists uniform illumination of the neon lamp across a wide area at the entrance slit to the spectrograph, the light at the input to the spectrograph can be modelled as a rectangle function. The effect on the image is once again a convolution, whereby the neon spectrum is convolved in one dimension with a *rect* function that approximates the slit:

$$\begin{aligned} Neon'(\lambda') &= Neon'(\lambda) * rect_W = \int Neon'(\lambda) rect_W(\lambda' = \lambda) d\lambda \\ rect_W(\lambda) &= 1 \ \forall \ |\lambda| < W/2 \end{aligned} \tag{7.12}$$

Parameter	Unit	Range of uniform distribution
Reflection(k)	NA	+1
Diffraction Order(n)	NA	-1
Grating Period (d)	lines/mm	1000 ± 5
Half the deviation angle (α)	degree	10.94 ± 2
Grating angle (θ_d)	degree	16.5 ± 9.5
Focal length (f)	mm	500 ± 1
Camera pixel pitch (T)	μm	26
Camera width (N)	pixels	1000
Camera centre position (C)	pixels	0
Slit width (W)	pixels	1-4
Neon line width (Γ)	nm	0.01-0.02

Table 7.2: The relationship between the pixel detectors and the wavelength axis for a Czerny-Tunrer spectrometer can be modelled using Equation 7.11. In order to simulate an arbitrary spectrometer, these parameters are randomised over a uniform distribution for the training set. The range of the uniform distribution is defined in the table. Also shown in the table is the slit which is allowed to vary continuously over a range of 1-4 times the width of the detector pixel, T. We note that narrowing the range of any of these variable would likely improve the accuracy of the trained networks.

In order to build in a variable slit into the simulation process, it is necessary to perform this convolution with the *rect* function of variable width W. Convolution can be implemented in a numerically efficient manner using the fast Fourier transform. This approach is based on the property of the Fourier transform, whereby the convolution of two functions in space is equivalent to the product of the Fourier transforms in the Fourier domain. Therefore, in order to compute Equation 7.12 we implement a DFT of the ideal neon function $Neon(\lambda)$, multiply by a discretised function (the Fourier transform of a function is a $rect$ function and vice versa) and performing an inverse DFT. In this way a variable slit can be built into the simulation model

7.3.4 Generating Training and Validation Datasets

The datasets used for training and validation were generated using Equation 7.10, Equation 7.11 and Equation 7.12. The first step is to randomly select the parameters that specify the physical spectrograph to be simulated. These parameters were slected from a uniform distribution as defined in Table 7.2. Once the parameters were selected, values of $x = -500 \rightarrow 499$ were input to Equation 7.11 in order to calculate the 'true wavelength axis' of length 1000 samples. Once the true wavelength axis has been calculated, the

neon spectrum can now be simulated for that spectrograph. The wavelength axis is upsampled by a factor of 10 and these wavelength values are input to Equation 7.10 to produce a neon spectrum of 10000 samples. In this equation the amplitude of the neon peaks is uniformly randomly varied between 90% to 110% from the relative intensity values in Table 7.1 and line width Γ is uniformly randomly varied between 0.01 nm to 0.02 nm for each peak in an attempt to capture the experimental variability in capturing a neon spectrum. For each simulated spectrum, normalization should be applied to make sure all the intensity is at the range of 0 to 1. The final step is application of a DFT to the 10000 long Neon spectrum followed by multiplication with the function (defined by width slit width W) and followed by an inverse DFT. The DFT operations were implemented using the fast Fourier transform. The next step is to downsample the output from the DFT to length of 1000 in order to match the size of the detector and the length of the original wavelength axis and the final step is the addition of a weak gaussian noise to simulate the effect of camera noise. Shot noise can be approximated as a Gaussian noise with high accuracy for high standard deviations. The upsampling/downsampling step is required by the DFT; in order to generate an accurate representation of the neon spectrum Fourier transform, a high sampling rate is required due to the the very narrow lineshape of the neon lines. The overall simulation process is implemented using the scipy python module.

Some examples of the simulated neon spectra that were generated using the above approach are shown in Fig. 7.4. It is clear significantly different wavelength axes will be used for each neon spectrum. The difference in the neon axes is more significant than a simple shifting, and could also be described by a non-linear warping.

In addition to the neon spectra, the LSTMs also require the wavelength axis for training. As described earlier, we do not train the LSTMs with true wavelength axis, rather they are trained to produce the four coefficients of the third order polynomial that approximates the true wavelength axis. These four coefficients are paired together with the matching neon spectrum as input to train the networks. It is important to emphasize that the third order polynomial fitting that is used here is not the same as the

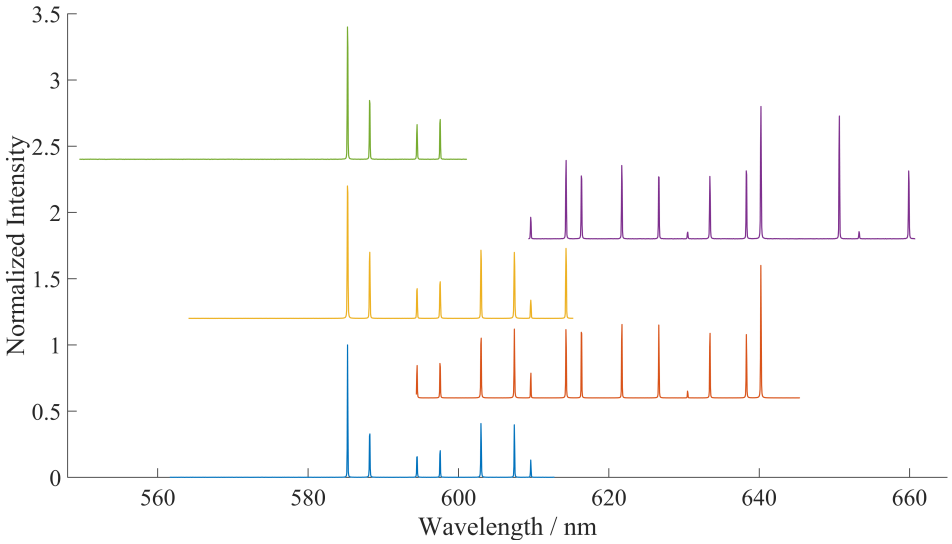


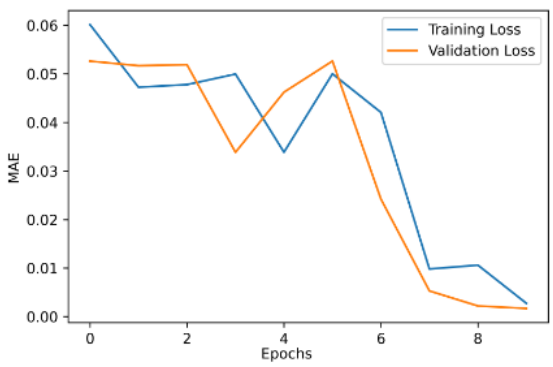
Figure 7.4: Examples of five different neon spectra that were simulated using the approach described in this chapter. Depending on the parameters for the spectrograph used in simulation, most notably the grating angle θ the spectra will be recorded from different wavelength bands.

traditional third order fitting used for wavelength calibration using some finite number of neon peaks. Here the third order fitting is applied to the 1000 samples that make up the true wavelength axis described earlier. As such it can be assumed to be significantly more accurate than traditional calibration polynomial fitting.

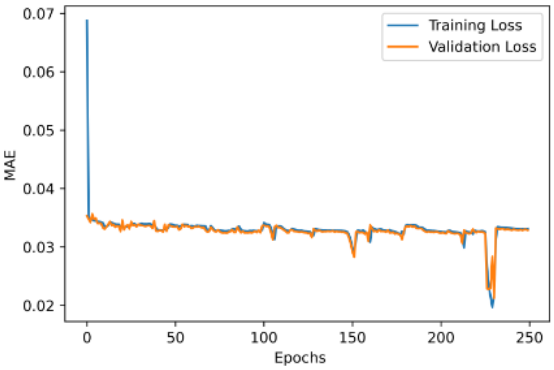
7.4 Results

In Fig. 7.5 the loss functions for the training and validation set are shown for the four encoder/decoder pairs that make up the 4-LSTM model. Different epoch numbers were used for the different cases as shown in the figure and further epochs did not improve any of the individual cases appreciably. Each encoder/decoder pair is trained to produce a single coefficient of the third-order polynomial that approximates the wavelength axis. In Fig. 7.6 the two loss functions for the training and validation sets are shown for the two encoder/decoder pairs that make of the 2-LSTM model. As for the previous case. Further epochs did not improve any of the individual cases appreciably. For the 2-LSTM model the first encoder/decoder pair is trained to produce a a_0, while the second encoder/decoder pair produces the other three coefficients that describe the third-order polynomial that approximates the wavelength axis. We note that the time taken to train the 2-LSTM model was less than half that required to train the 4-LSTM model case and there is only a marginal improvement as described below for the case of the 4-LSTM model.

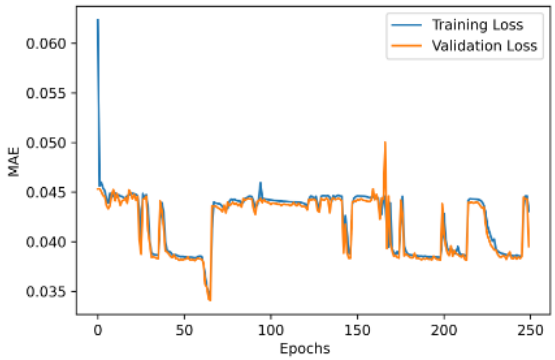
In Fig. 7.7, Fig. 7.8, and Fig. 7.9 results are shown for three different reference spectra that are calibrated with the trained 4-LSTM model. These three cases differ in terms of the number of neon peaks that are contained within the reference spectrum, and in each case the results are found to differ appreciably. In Fig. 7.7 the result of applying the trained 4 LSTM model is shown for the case of a neon spectrum containing only five neon peaks shown in Fig. 7.7 (a). In Fig. 7.7 (b) the true wavelength axis as a function of detector pixel position is plotted, as well as the wavelength axes predicted by the 4-LSTM model and traditional third order polynomial fitting applied to the five neon peaks. The MAE for the 4-LSTM result is 0.04811nm and for traditional third order fitting it is 0.11431nm. Three regions are highlighted in the figure and magnified in Fig. 7.7 (c), (d) and (e) in which it can be seen that the accuracy of the two methods varies over the range of the detector. Interestingly, it can be seen that a spectrum with few peaks that are condensed on one side of the spectrum, the accuracy of the 4-LSTM model is more accurate than third order polynomial fitting in regions that are far away from the



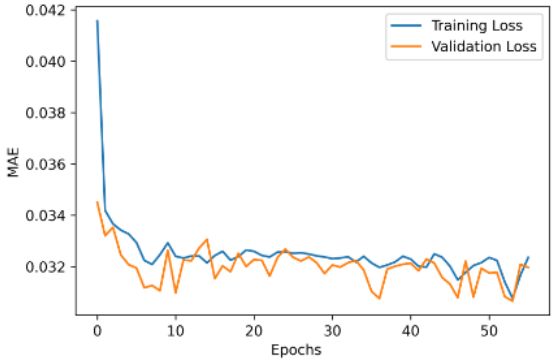
(a) LSTM Encoder/Decoder 1 - Relates to output a_0



(b) LSTM Encoder/Decoder 2 - Relates to output a_1

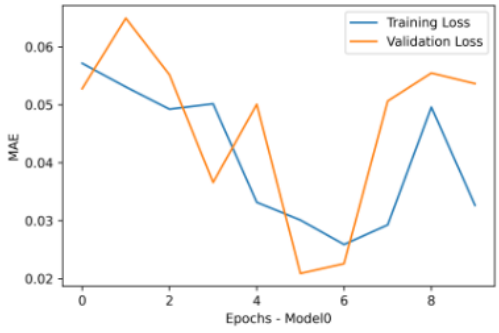


(c) LSTM Encoder/Decoder 3 - Relates to output a_2

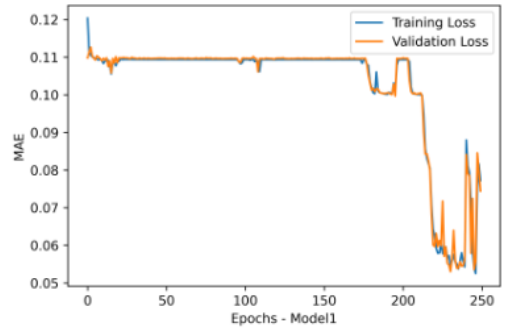


(d) LSTM Encoder/Decoder - Relates to output a_3

Figure 7.5: Loss functions for training and validation sets used over various epoch numbers for the four encoder/decoder pairs that make use of the 4-LSTM model. Further epochs did not improve any of the individual cases appreciably.



(a) LSTM Encoder/Decoder 1 - Relates to output a_0



(b) LSTM Encoder/Decoder 2 - Relates to output a_1, a_2, a_3

Figure 7.6: Loss functions for validation sets used over various epoch numbers for the four encoder/decoder pairs that make use of the 2-LSTM model. Further epochs did not improve any of the individual cases appreciably.

reference peaks, while within the region of the peaks, third order fitting is more accurate. In order to elucidate this point further the wavelength error of both methods is plotted in Fig. 7.7 (d). Overall the 4-LSTM model performs better than 3rd order fitting over the full range. A similar set of results is shown in Fig. 7.8 for the case of a neon spectrum containing seven peaks, as shown in Fig. 7.8 (a). In this case the MAE for the 4-LSTM network is 0.07067 nm. In Fig. 7.8 (b) the true wavelength axis is shown together with the wavelength axes predicted by the 4-LSTM model and traditional third order polynomial fitting. The three regions are highlighted in the figure and magnified in Fig. 7.8 (c), (d) and (e) in which it can be seen that the accuracy of the 4-LSTM model appears to reduce with distance from the region in which the peaks are condensed but the 3rd order fitting method is more accurate over the full range. We continue this analysis with a neon spectrum containing 12 neon peaks in Fig. 7.9(a). The result of inputting this spectrum to the 4-LSTM model is shown in Fig. 7.9(a) together with the true wavelength axis and the result of traditional third order polynomial fitting applied to the twelve neon peaks. The MAE for the 4-LSTM network is 0.02542 nm and for traditional third order fitting is 0.00134 nm. The wavelength error of both methods is plotted in (d) and it is clear that 3rd order fitting is significantly more accurate than the 4-LSTM model.

A matching set of results are shown in Fig. 7.10, Fig. 7.11, and Fig. 7.12 and the same set of conclusions can be drawn: both the 2-LSTM model and the 4-LSTM model can provide higher wavelength calibration accuracy than traditional third order fitting of the neon peaks, if there are five or less peaks in the spectrum. Similar accuracy is obtained for six peaks and as the number of peaks increases, the accuracy of third order fitting improves, while there is no improvement in either LSTM approach. In order to further examine dependence of accuracy on peak number, three tests sets were generated of size 1000 and passed as input to the 4-LSTM network and another set of three datasets were passed to the 2-LSTM network. The same parameters were used to model the spectrograph in the simulation process to generate these six test sets as defined in Table 7.2 except for case of the grating angle, which was varied over three different ranges as defined in this table. The effect of using these ranges is to

Model Type	Theta(degree)	LTSM(nm)	Trad3(nm)
2-LSTM Model	16.5±0.9	0.014186	0.006438
	16.5±1	0.044101	0.015023
	16.5±1.1	0.045389	0.07701
4-LSTM Model	16.5±0.9	0.044067	0.006114
	16.5±1	0.043551	0.029205
	16.5±1.1	0.043171	0.046269

Table 7.3: This table shows the MAE results for three different test sets that were applied to the trained network. For these test sets, the parameters of the spectrograph were varied according to the values given in Table 7.2 except for case of the grating angle, which was varied over three different ranges as defined in this table. The effect of using these ranges is to control approximately the number of neon peaks that will appear in the window of the spectrum. The first test set with the smallest range of θ will have the largest number of peaks on average, and this number will drop statistically across the test set as the range of θ is increased. The result is that the accuracy of third order polynomial fitting drops significantly as the test sets can contain spectra with fewer peaks, while the two LSTM models provide more consistent results and outperform third order fitting for the less populated test sets. It is notable that the 4-LSTM model is three times more accurate than the 2-LSTM model for the most dense test set, but the accuracy is similar for the lesser populated ones. The performance of the two LSTM models for reference spectra with fewer peaks is further examined in Table 7.4

control approximately the number of neon peaks that will appear in the window of the spectrum. The first test set with the smallest range of θ will have the largest number of peaks on average, and this number will drop statistically across the test set as the range of θ is increased. The result is that the accuracy of third order polynomial fitting drops significantly as the test sets can contain spectra with fewer peaks, while the two LSTM models provide more consistent results and outperform third order fitting for the less populated test sets. It is notable that the 4-LSTM model is three times more accurate than the 2-LSTM model for the most dense test set, but the accuracy is similar for the lesser populated ones. The performance of the two LSTM models for reference spectra with fewer peaks is further examined in Table 7.4, which shows the MAE results for several test sets that were carefully controlled to produce the same number of peaks for all spectra in the test set, and subsequently applied to the trained network. To achieve this fine-grained control, the parameters of the spectrograph were fixed as indicated in the left column except for the case of the grating angle, which was varied over different ranges as defined in this table in order to provide some degree of variability across the

1000 spectra in the test sets. The number of peaks shown in the test sets are shown in the right most column. It is clear that third order polynomial fitting has superior accuracy than both LSTM models when the reference spectrum contains six peaks or more but the superiority of third order fitting is clearly less pronounced as the number of peaks reduces and two LSTM models out-perform third order fitting for five peaks and lower. The 4-LSTM model performs only marginally better than the 2-LSTM model over all cases. The values in this table are presented graphically in two bar charts in Fig. 7.13 and Fig. 7.14 and it is clear that the 4-LSTM model outperforms third order fitting when there are only five peaks or less in the neon spectrum for both cases.

Model type	Theta range (min,max)		LTSM(nm)	Trad3(nm)	Peak Number
Model 2 (f=0.499m, alpha=8.94, d=995 l/mm, c=-10)	15.4	15.53	0.121998	0.189281	4
	15.57	15.66	0.078472	0.102979	5
	15.7	15.73	0.059564	0.022343	6
	15.77	15.87	0.044975	0.016788	7
	15.91	15.93	0.035537	0.005322	8
	15.97	16.1	0.032531	0.006146	9
	16.14	16.25	0.025616	0.002172	10
	16.29	16.36	0.025983	0.001137	11
	16.4	16.45	0.029315	0.001166	12
Model 4 (f=0.499m, alpha=8.94, d=995 l/mm, c=-10)	15.4	15.53	0.118675	0.171513	4
	15.57	15.66	0.07807	0.121962	5
	15.7	15.73	0.057178	0.021853	6
	15.77	15.87	0.046206	0.017122	7
	15.91	15.93	0.037208	0.004854	8
	15.97	16.1	0.029686	0.00506	9
	16.14	16.25	0.025398	0.001782	10
	16.29	16.36	0.024269	0.00134	11
	16.4	16.45	0.025753	0.00127	12

Table 7.4: This table shows the MAE results for several test sets that were carefully controlled to produce the same number of peaks for all spectra in the test set, and subsequently applied to the trained network. To achieve this fine-grained control, the parameters of the spectrograph were fixed as indicated in the left column except for the case of the grating angle, which was varied over different ranges as defined in this table in order to provide some degree of variability across the 1000 spectra in the test sets. The number of peaks shown in the test sets is shown in the right most column. It is clear that third order polynomial fitting has superior accuracy than both LSTM models when the reference spectrum contains six peaks or more but the superiority of third order fitting is clearly less pronounced as the number of peaks reduces and two LSTM models out-perform third order fitting for five peaks and lower. The 4-model LSTM model performs only marginally better than the 2-LSTM model over all cases. The values in this table are presented graphically in two bar charts in Fig. 7.13 and Fig. 7.14

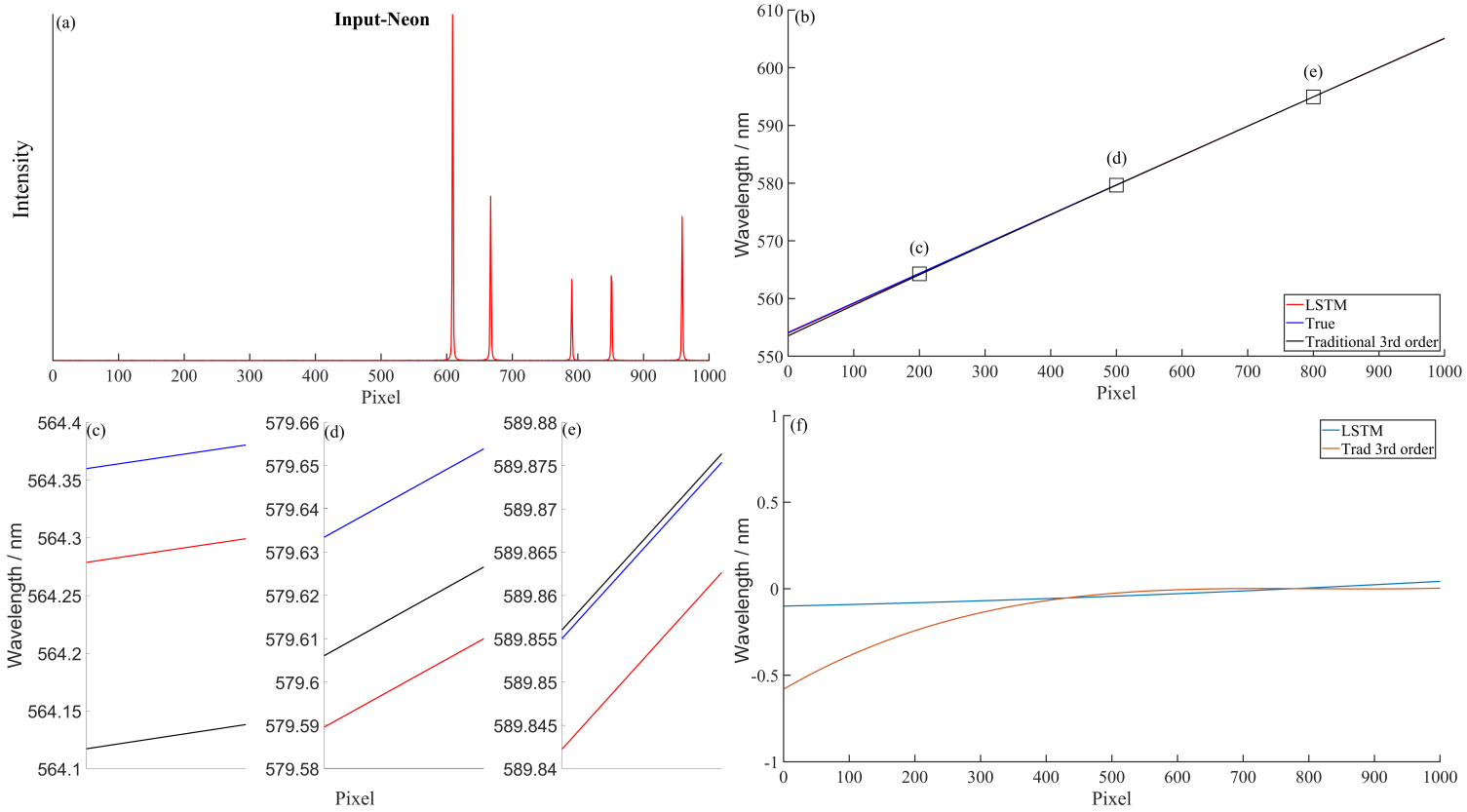


Figure 7.7: 4 LSTM model for the case of only 5 neon peaks: The MAE for 4-LSTM is 0.04811nm and for traditional third order fitting is 0.11431nm. (a) the neon spectrum with 5 peaks that was passed as input to the 4-LSTM model; (b) shows the true wavelength axis as a function of detector pixel position, as well as the wavelength axes predicted by the 4-LSTM model and traditional third order polynomial fitting applied to the five neon peaks. Three regions are highlighted in the figure and magnified in (c), (d) and (e) in which it can be seen that the accuracy of the two methods varies over the range of the detector. Interestingly it can be seen that a spectrum with few peaks that are condensed on one side of the spectrum, the accuracy of the 4-LSTM model is more accurate that 3rd-order polynomial fitting in regions that are far away from the reference peaks, while within the region of the peaks third order fitting is more accurate. In order to elucidate this point further the wavelength error of both methods is plotted in (d). Overall the 4-LSTM model performs better than 3rd order fitting over the full range.

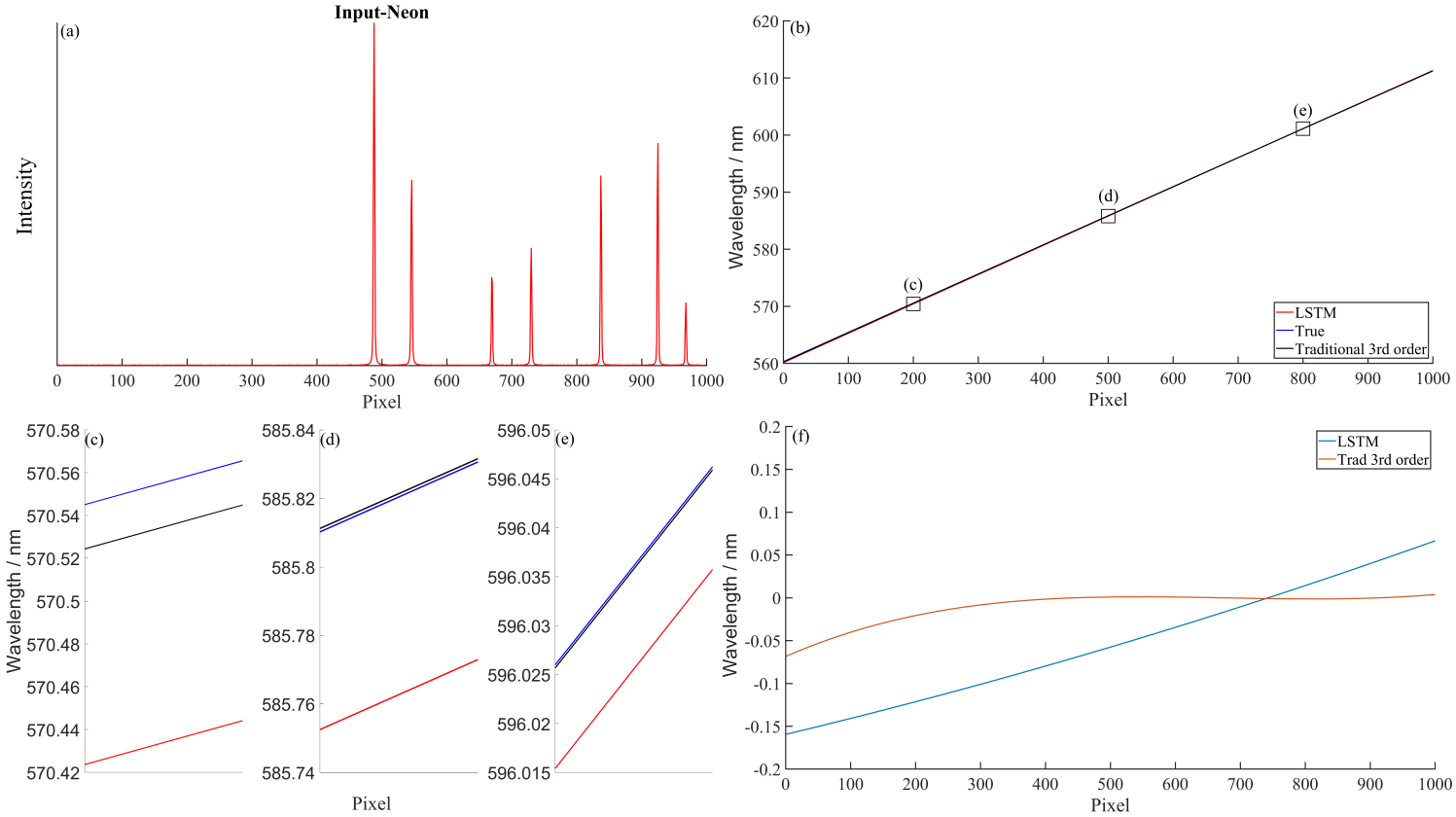


Figure 7.8: 4-LSTM model for the case of 7 neon peaks: The MAE for 4-LSTM is 0.07067 nm and for traditional third order fitting is 0.07067 nm. (a) the neon spectrum with 7 peaks that was passed as input to the 4-LSTM model; (b) shows the true wavelength axis as a function of detector pixel position, as well as the wavelength axes predicted by the 4-LSTM model and traditional third order polynomial fitting applied to the five neon peaks. Three regions are highlighted in the figure and magnified in (c), (d) and (e) in which it can be seen that the accuracy of the two methods varies over the range of the detector. In this case the accuracy of the 4-LSTM model is less accurate that 3rd-order polynomial fitting in all three regions regardless of distance from the reference peaks. The wavelength error of both methods is plotted in (d). It can be seen that the accuracy of the 4-LSTM model reduces with distance from the region in which the peaks are condensed but the 3rd-order fitting method is more accurate over the full range.

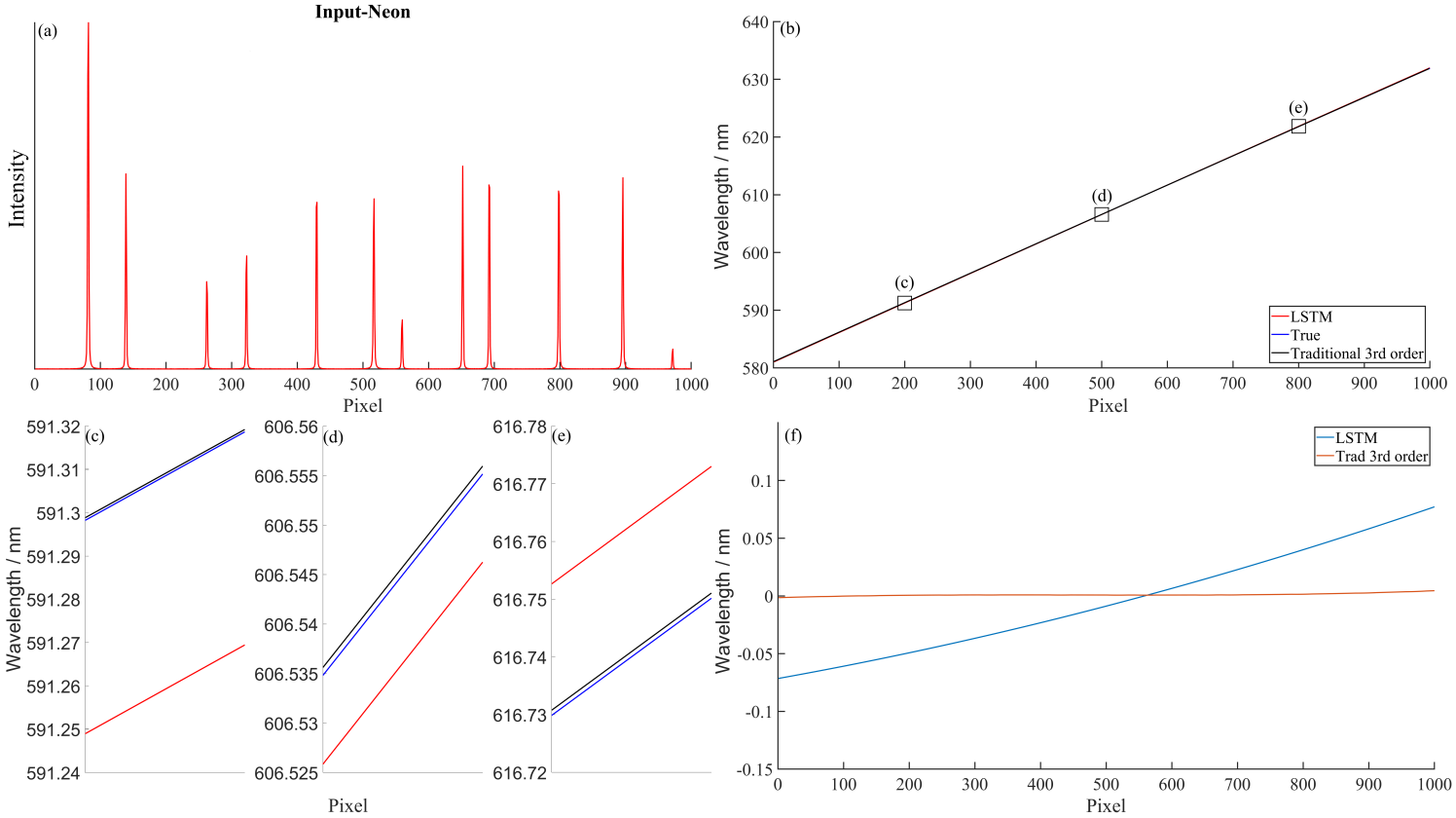


Figure 7.9: 4 LSTM model for the case of 12 neon peaks: The MAE for 4-LSTM is 0.02542 nm and for traditional third order fitting is 0.00134 nm. (a) the neon spectrum with 12 peaks that was passed as input to the 4-LSTM model; (b) shows the true wavelength axis as a function of detector pixel position, as well as the wavelength axes predicted by the 4-LSTM model and traditional third-order polynomial fitting applied to the five neon peaks. Three regions are highlighted in the figure and magnified in (c), (d) and (e) in which it can be seen that the accuracy of the 4-LSTM model is once again less accurate than 3rd-order polynomial fitting in all three regions and the difference in accuracy is more pronounced. The wavelength error of both methods is plotted in (d).

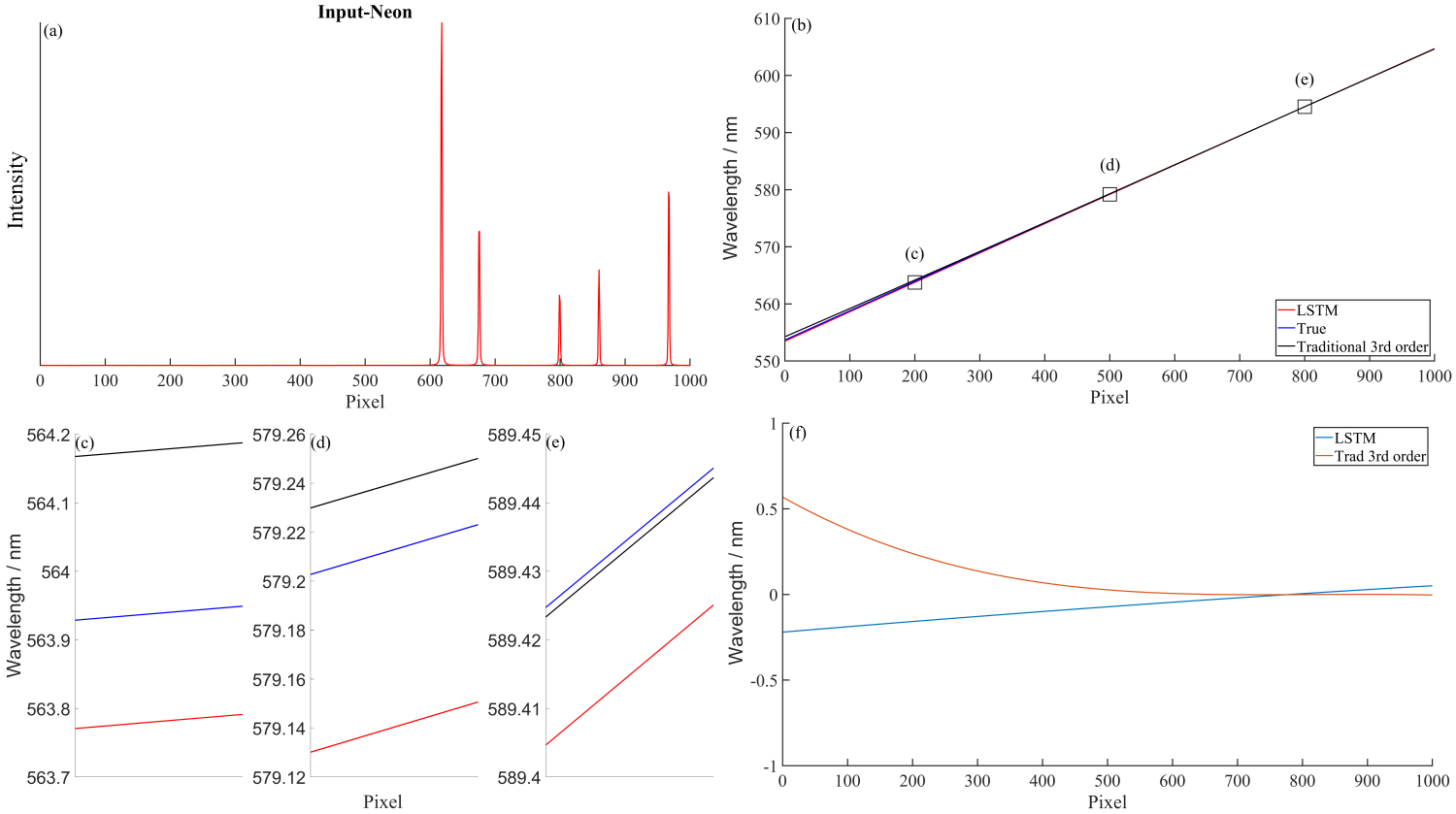


Figure 7.10: 2 LSTM model for the case of only 5 neon peaks: The MAE for 4-LSTM is 0.08763 nm and for traditional third order fitting is 0.11242 nm. (a) the neon spectrum with 5 peaks that was passed as input to the 4-LSTM model; (b) shows the true wavelength axis as a function of detector pixel position, as well as the wavelength axes predicted by the 2-LSTM model and traditional third order polynomial fitting applied to the five neon peaks. Three regions are highlighted in the figure and magnified in (c), (d) and (e) in which it can be seen that the accuracy of the two methods varies over the range of the detector. Interestingly it can be seen that a spectrum with few peaks that are condensed on one side of the spectrum, the accuracy of the 2-LSTM model is more accurate that 3rd order polynomial fitting in regions that are far away from the reference peaks, while within the region of the peaks third order fitting is more accurate. In order to elucidate this point further the wavelength error of both methods is plotted in (d). Overall the 4-LSTM model performs better than 3rd order fitting over the full range.

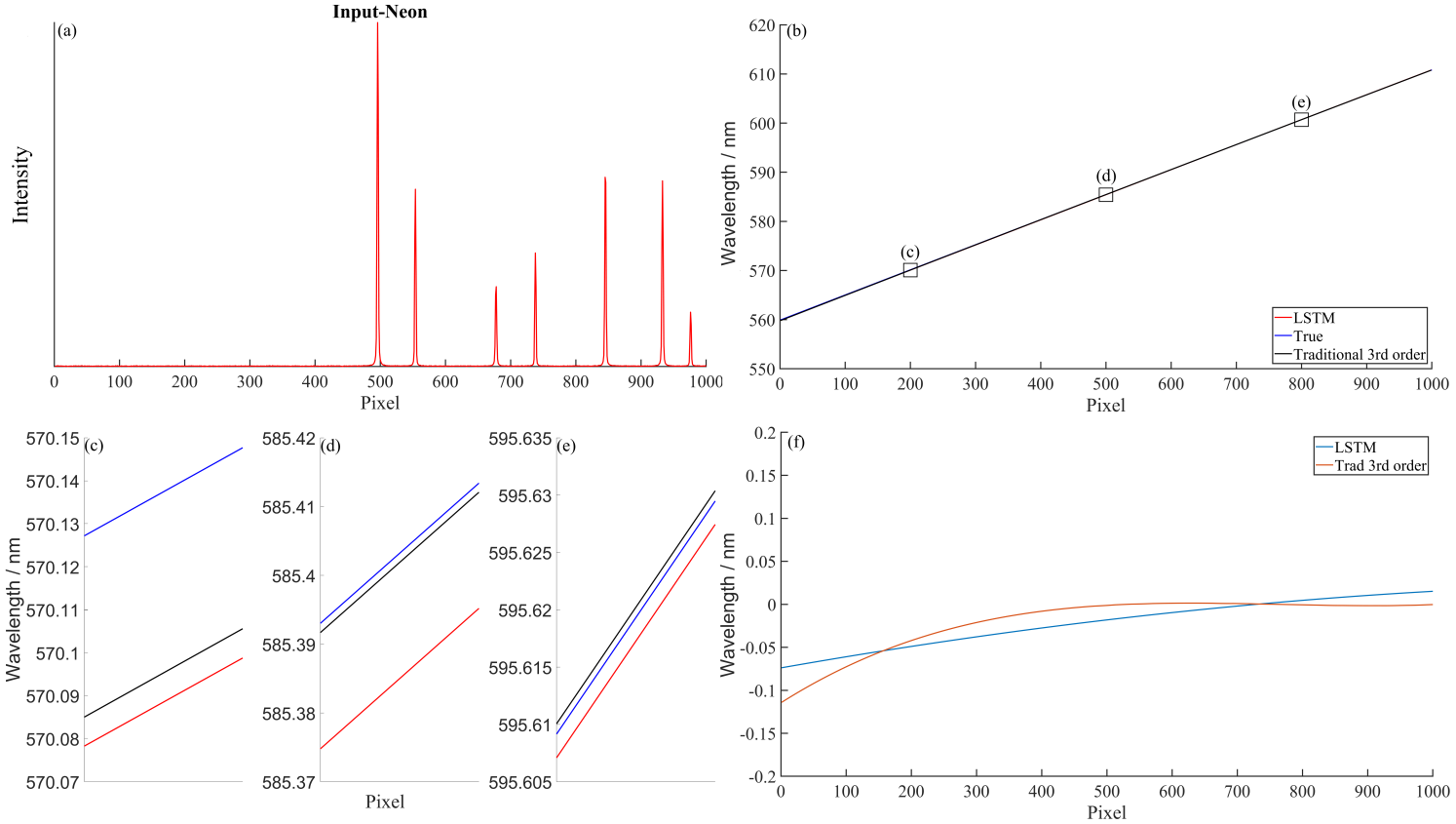


Figure 7.11: 2 LSTM model for the case of 7 neon peaks: The MAE for 4-LSTM is 0.02627 nm and for traditional third order fitting is 0.02025 nm. (a) the neon spectrum with 7 peaks that was passed as input to the 2-LSTM model; (b) shows the true wavelength axis as a function of detector pixel position, as well as the wavelength axes predicted by the 2-LSTM model and traditional third order polynomial fitting applied to the five neon peaks. Three regions are highlighted in the figure and magnified in (c), (d) and (e) in which it can be seen that the accuracy of the two methods varies over the range of the detector. The wavelength error of both methods is plotted in (d) over the full range. It can be seen that the accuracy of both methods reduces with distance from the region in which the peaks are condensed and both methods have similar accuracy over the full range.

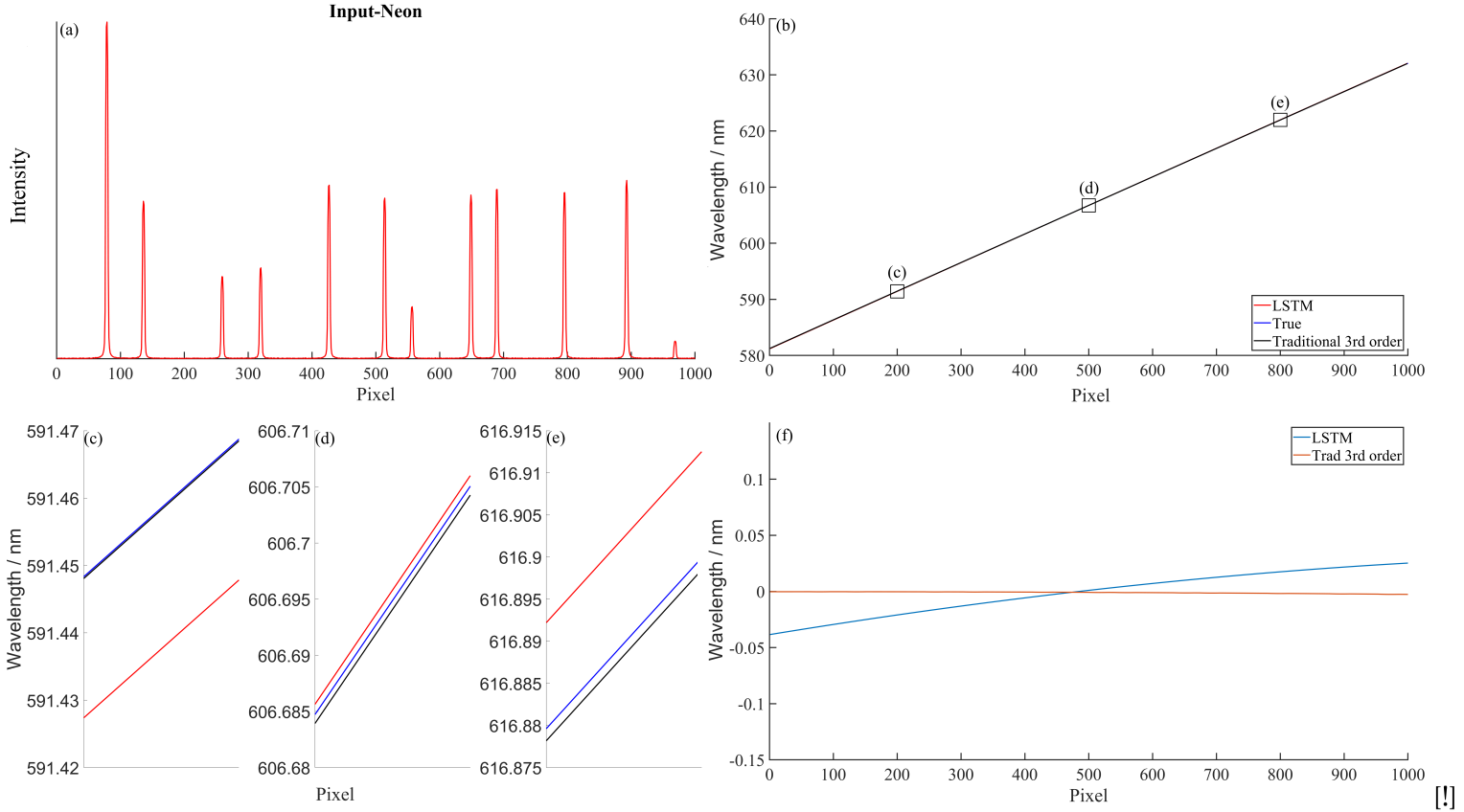


Figure 7.12: 2-LSTM model for the case of 12 neon peaks: The MAE for 4-LSTM is 0.016 nm and for traditional third order fitting is 0.0001027 nm. (a) the neon spectrum with 12 peaks that was passed as input to the 2-LSTM model; (b) shows the true wavelength axis as a function of detector pixel position, as well as the wavelength axes predicted by the 2-LSTM model and traditional third order polynomial fitting applied to the five neon peaks. Three regions are highlighted in the figure and magnified in (c), (d) and (e) in which it can be seen that the accuracy of the 2-LSTM model has the same level of accuracy as 3rd order polynomial fitting in the centre region but this reduces with distance from centre. The wavelength error of both methods is plotted in (d).

[!]

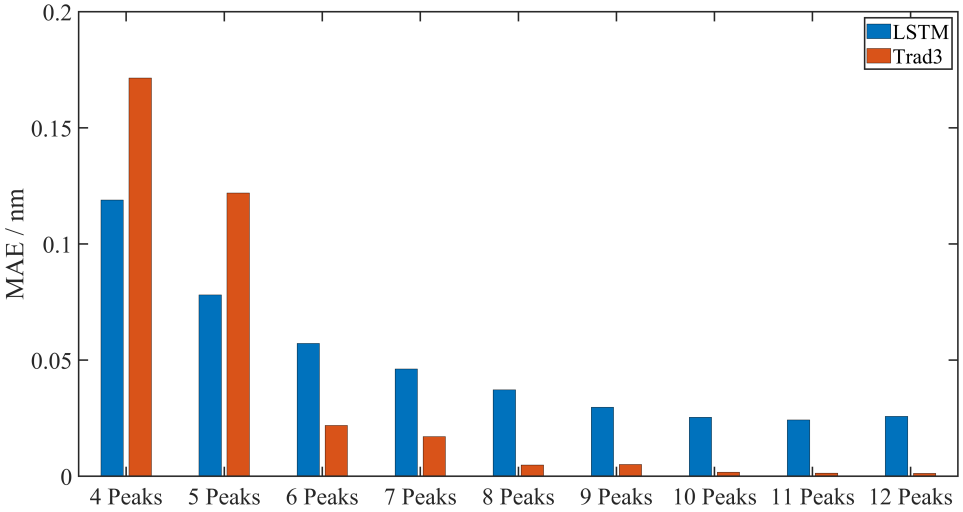


Figure 7.13: Comparison of MAE results for the 4-LSTM model when applied to test sets containing different numbers of peaks. These values are related to those provided in Table 7.4. The 4-LSTM model outperforms third order fitting when there are only five peaks or less in the neon spectrum.

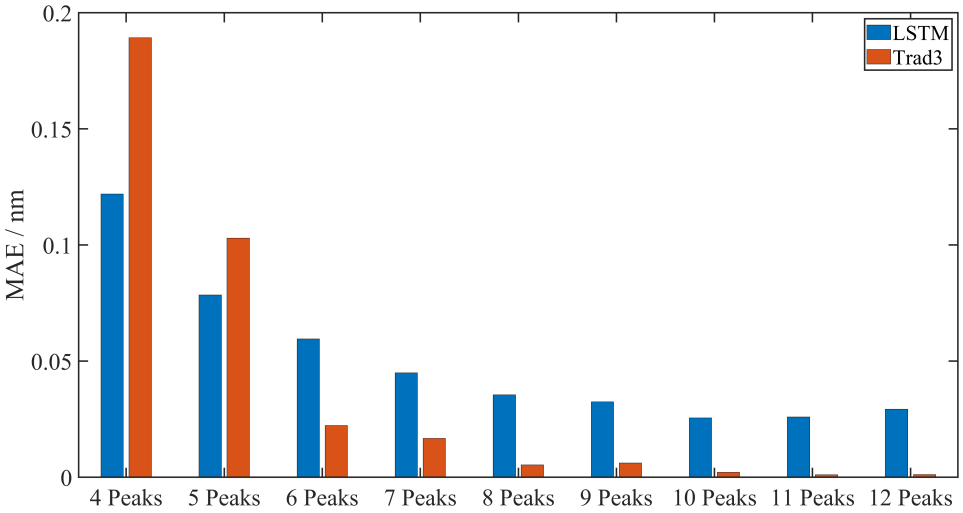


Figure 7.14: Comparison of MAE results for the 2-LSTM model when applied to test sets containing different numbers of peaks. These values are related to those provided in Table 7.4. The 2-LSTM model outperforms third order fitting when there are only five peaks or less in the neon spectrum. The results for the 2-LSTM model are similar to those for the 4-LSTM model.

7.5 Conclusion

In this chapter, we have investigated the use of the long term short term neural network (LSTM) for the wavelength calibration of spectrometers. LSTMs are a versatile machine learning tools that are a sub class of the broader recurrent neural network. LSTMs are particularly useful in learning time sequence patterns that recur, but do so at irregular time intervals. In particular we use the encoder/decoder LSTM architecture as developed by Google, which facilitates the use of different lengths for the input and output sequences. Like all deep learning methods, training with many thousands of examples is required. Here we train with neon spectra and the matching wavelength axis in the form of four coefficients that describe the 3rd order polynomial that (accurately) approximates the true wavelength axis. It is important to emphasise that this is not the same as traditional third order polynomial fitting used for wavelength calibration, which makes use for only a small number of peaks. Here we use the entire wavelength axis at every point for fitting.

A key contribution in this chapter is the development of a simulation method capable of generating thousands of simulated neon spectra. This simulation is based on modelling a Raman spectrometer in terms of the spectrograph parameters, including the focal length, the grating angles, slit width, and the detector specifications. The simulator can produce hundreds of thousands of simulated neon spectra in a short time, together with the correct (non-linear) wavelength axis, which can then be used for training with machine learning. It is possible that other researchers will make further use of this simulator as better methods in deep learning are developed for this application.

This chapter marks the first effort in using deep learning to directly calibrate a spectrometer using a reference atomic emission spectrum. These initial results indicate that the approach can provide superior results when compared with polynomial fitting if there are a low number of peaks within the bandwidth of the spectrometer. However, as the number of peaks increases, polynomial fitting outperforms the proposed method. We believe that further work may show that machine learning can outperform polynomial fitting in all cases.

For the case of Raman spectroscopy, several wavenumber reference standards are also available for use with polynomial fitting, such as 4-acetamidophenol or benzonitrole [3]. Raman spectra can be calibrated using polynomial fitting of either a wavenumber reference or a wavelength reference standard, with the latter requiring the additional step of acquiring an accurate measurement of the laser wavelength. We believe that the approach proposed in this chapter could also be applied to wavenumber reference standards. This would require the simulator to include the conversion of the wavelength to wavenumber and would also require the use of the laser wavelength. This offers another possible avenue of research for this work.

When it comes to machine learning models, repeatability and generality are important factors to consider. Repeatability refers to the ability of a model to produce consistent and reproducible results when presented with the same input data multiple times. This is crucial for ensuring that the model can be relied upon for consistent predictions and decision making. To achieve repeatability, it is important to use consistent data preprocessing and feature engineering methods, as well as to tune the hyperparameters of the model carefully. Generality refers to the ability of a model to generalize well to new, unseen data that was not used during training. This is important for ensuring that the model can be applied to real-world scenarios and is not simply overfitting to the training data. To achieve generality, it is important to use a diverse range of training data that is representative of the real-world data that the model is intended to be used on. Additionally, regularization techniques such as dropout or weight decay can be used to prevent overfitting. It is true that, in some cases, training a machine learning model can be seen as simply trading calibrating a machine for calibrating a machine learning model. However, there are important differences between the two approaches. For example, machine learning models can potentially learn more complex relationships between inputs and outputs than traditional calibration methods, allowing for more accurate and precise predictions. Additionally, machine learning models can be updated and retrained with new data, allowing them to continually improve over time.

However, it is important to note that machine learning models can also be subject

to biases and limitations, and may not always be appropriate for every application. It is important to carefully evaluate the performance and limitations of any machine learning model before applying it in a real-world setting.

Chapter 8

Conclusion

Raman spectroscopy can probe the chemical structure of a material. The uniqueness of the spectrum from various materials has resulted in the term 'fingerprint' being used to describe the Raman spectrum. Indeed, such is the capacity of Raman spectroscopy to identify different materials, it can be used to classify biological cells and tissue as they change from a healthy state into a diseased state. Such a form of diagnosis is sometimes called an 'optical biopsy' and various research groups have proved the potential of the technique for a range of different cancers. Recent progress includes the mounting of Raman probes onto hypodermic needles and automated cytology systems that can probe individual cells taken non-invasively from the mouth, bladder or cervix. The Raman spectrum from a healthy cell and diseased cells are subtly different. A key component in Raman diagnostics is the use of multivariate statistical algorithms that can be trained using datasets of known samples to classify the groups based on the subtle differences between them.

Despite the many hundreds of millions of euros, the thousands of journal paper proving the usefulness of Raman diagnostics, and the significant reduction in cost of Raman spectrometers in recent decades, Raman spectroscopy has never been adopted clinically. The key reason for this is the poor resproducibility of Raman spectroscopy across instruments; in other words the same material can produce different spectra when recorded using different spectrometers. These differences can include small movement of the Raman peaks along the wavenumber axis (wavenumber miscalibration) or

random changes in the amplitude of the peaks (intensity calibration). Such changes can render a multivariate classifier trained on one instrument to be completely useless in identifying samples recorded from another instrument. This poor cross-instrument comparability has been highlighted by two recent multi-site studies involving 26 and 35 instruments respectively, both of which demonstrated inconsistencies in the wavenumber shift for various materials even following established calibration protocols provided by the instrument manufacturer. There is clearly a growing need for a set of calibration protocols, that are inexpensive, accessible to, and easy to implement by, all Raman researchers, and above all are accurate enough to ensure reliable cross-instrument comparability.

The overall goal of this thesis is to develop a new method for wavenumber and intensity calibration that can help Raman spectroscopy penetrate into the clinic. To this end several contributions have been made. Firstly the reviews of wavelength and wavenumber calibration in Chapter 2 can be considered valuable contributions. Secondly, The core topic of Chapters 4 and 7 is the development of new methods of wavelength calibration, often a precursor of both wavenumber and intensity calibration in Raman spectroscopy. The approach proposed in Chapter 4 is to develop an algorithm that takes in a neon spectrum as input and outputs the true wavelength axis; this algorithm is based on a **model** that relates the wavelength and the detector pixels using the physics of optical imaging and diffraction and provides greater accuracy than traditional methods, especially in bands where there are few neon lines. The approach in Chapter 7 is to use the same **model** build a simulator that can generate hundreds of thousands of fake neon spectra (with their matched true wavelength axis) to train a machine learning network based on the recently proposed LSTM, which is designed to learn patterns in 1D sequences of data. This method does not provide greater accuracy than traditional methods except when there are five or fewer neon lines in the spectrum, which can happen experimentally for some spectrometer configurations. In this proof-of-concept study we examined a very basic 'off-the-shelf' LSTM architecture and we believe that better results could be obtained with further development of the idea and further ex-

ploration of more complex machine learning architectures. The method could also be extended to work for direct wavenumber calibration using a reference material like benzonitrile or similar; this would require an extension of the simulator in the same way that Chapter 5 extends the physical model developed in Chapter 4.

While wavelength calibration is the subject of Chapters 4 and 7, direct wavenumber calibration is the subject of Chapter 3 and 5. Direct wave calibration makes use of a reference Raman material containing several sharp spectral lines such an benzonitrile or 4-acetamidophenol and is preferable over the wavelength/conversion approach as it is single step and does not require measurement of the laser wavelength while providing a similar level of accuracy. In Chapter 3 we investigate a new Raman wavenumber reference material in the form of a commercial plastic, which has several advantages over commmonly used reference materials in terms of low cost, simple mounting, and high photo-stability. Although the accuracy afforded by the polymer does not match that of 4-acetamidophenol or benzonitrile, the accuracy is sufficient for many applications and may be an attractive option in many cases. There is also the consideration that this material could also be used as part of an intensity calibration routine as described in Chapter 6, and might therefore provide a dual use. The work in Chapter 5 is an extension of Chapter 4, whereby the physical model that relates wavelength to detector pixel for an arbitrary spectrometer is extended to relate wavenumber to pixel for an arbitrary Raman spectrometer. The resulting algorithm is demonstrated on hundreds of spectra to provide greater accuracy than traditional polynomial fitting in general, and especially in bands where there are few spectral lines.

The remaining chapter, Chapter 6, is the only chapter to focus on intensity calibration. Although wavenumber calibration is commonly applied by Raman researchers, intensity calibration if far less commonly applied in the literature. We believe there are several reasons for this: firstly wavenumber calibration is always required if there is any attempt is to be made on wavenumber assignment; for example a sharp peak in a biological spectrum at 1004 cm^{-1} is known to relate to phenylalanine. However, most research papers use only a single system to record results and cross instrument

comparability is not a concern; therefore, intensity calibration is deemed superfluous to requirements since it requires expensive equipment that is often difficult to mount or align. The consequence is that datasets that are acquired over these hundreds of studies cannot be easily shared or combined. If a simple, inexpensive, and accurate method existed for intensity calibration it would be far more likely all researchers would adopt it and the overall research area would be far better off. Even those studies that have attempted to demonstrate cross-instrument comparability and have employed intensity calibration have shown that the results are not ideal. In most cases of intensity calibration in the literature, a NIST-calibrated white light lamp is preferred, which appears to be far more popular than the NIST fluorescent materials that must be purchased separately (and expensively) for each excitation wavelength and are tricky to align. However, these lamps must be frequently re-calibrated which is both expensive and time consuming and the laboratory may be without the lamp for several months while it sent away for re-calibration. In Chapter 6 we demonstrated how an arbitrary un-calibrated tungsten halogen lamp can be used to achieve higher accuracy than a NIST-calibrated lamp. Since most Raman microscopes are built around an existing Olympus microscope or similar, a tungsten halogen lamp is readily available and there is no reason to mount any sample holder or fiber to deliver the light. Furthermore, there is no need to worry about miscalibration of the lamp, which is obviously a factor for NIST-calibrated sources. Indeed, this latter point may explain why the results for our approach outperform those for a NIST calibrated lamp is most cases. There exists a secondary contribution in Chapter 6 that should not be overlooked; in it we propose a novel metric that can be used to quantify the accuracy of an intensity calibration protocol across two more instruments which takes into account the variable background that can occur in Raman spectroscopy. Previous attempts to gauge the performance of an intensity calibration protocol have been qualitative in nature.

If there was one subject that could be chosen by this author for future work it would be the intensity calibration protocol in Chapter 6. There is little doubt that intensity calibration is the single most significant problem hindering clinical adoption of Raman

today. Further experimental evidence is required to conclusively demonstrate that our method could become a standard across the Raman community. One limitation in Chapter 6 is that the spectra were limited to the fingerprint region, while spectra that cover the full band up to 4000cm^{-1} will likely suffer from the greatest need for sensitivity correction. Secondly, Chapter 6 did not consider the effect of polarisation and more accurate results can be expected by both NIST-calibrated white light correction and our own arbitrary white light correction if polarisation is taken into account.

A final note in this conclusion relates to the concensus standards that already exists for Raman instrumentation published by the American Society for Testing and Materials (ASTM) international in relation to performance testing, calibration, and relative intensity correction (ASTM E1683 [46], E1840 [2], E2529 [47], E2911 [48]). In the context of this thesis, several of these standards are relevant. For example, Chapter 3, 4, 5, and 7 relate to new methods for wavelength/wavenumber calibration in Raman spectroscopy, and for these chapters, the most relevant of the current set of standards is ASTM-E1840, [2] most recently updated in 2013, which focuses on Raman shift (or wavenumber) calibration. In the context of Chapter 6, the most relevant of the existing standards documents is ASTM-E2911, [48] also most recently updated in 2013, which focuses on Raman intensity calibration using the NIST Standard Reference Materials (SRMs) in the 224X series. It is my great hope that when these two standards documents are next revised, one or more of the methods developed in this thesis will be adopted.

References

[1] A. Kramida, Yu. Ralchenko, J. Reader, and and NIST ASD Team. NIST Atomic Spectra Database (ver. 5.7.1), [Online]. Available: `https://physics.nist.gov/asd` [2020, July 28]. National Institute of Standards and Technology, Gaithersburg, MD., 2019.

[2] ASTM E1840-96(2014), "Standard guide for Raman Shift standards for spectrometer calibration," standard, ASTM International, West Conshohocken,Pennsylvania, 2014.

[3] D. Liu, H. J. Byrne, L. O'Neill, and B. Hennelly, "Investigation of wavenumber calibration for raman spectroscopy using a polymer reference," in *Optical Sensing and Detection V*, vol. 10680, pp. 486–497, SPIE, 2018.

[4] D. A. Long, "Raman spectroscopy," *New York*, vol. 1, 1977.

[5] D. A. Long and D. Long, *Raman spectroscopy*. New York: McGraw-Hill, 1977.

[6] J. D. Santillán, C. D. Brown, and W. Jalenak, "Advances in raman spectroscopy for explosive identification in aviation security," in *Optics and Photonics in Global Homeland Security III*, vol. 6540, International Society for Optics and Photonics, 2007.

[7] D. I. Ellis and R. Goodacre, "Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and raman spectroscopy," *Analyst*, vol. 131, no. 8, p. 875–885, 2006.

[8] M. Diem, "Molecular pathology via ir and raman spectral imaging," *Journal of biophotonics*, vol. 6, no. 11-12, p. 855–886, 2013.

[9] G. Clemens, "Vibrational spectroscopic methods for cytology and cellular research," *Analyst*, vol. 139, no. 18, p. 4411–4444, 2014.

[10] L. Kerr and B. Hennelly, "A multivariate statistical investigation of background subtraction algorithms for raman spectra of cytology samples recorded on glass slides," *Chemometrics and Intelligent Laboratory Systems*, vol. 158, p. 61–68, 2016.

[11] L. T. Kerr, H. J. Byrne, and B. M. Hennelly, "Optimal choice of sample substrate and laser wavelength for raman spectroscopic analysis of biological specimen," *Analytical Methods*, vol. 7, no. 12, p. 5041–5052, 2015.

[12] K. Kong, C. Kendall, N. Stone, and I. Notingher, "Raman spectroscopy for medical diagnostics—from in-vitro biofluid assays to in-vivo cancer detection," *Advanced drug delivery reviews*, vol. 89, pp. 121–134, 2015.

[13] J. Hutchings, C. Kendall, B. Smith, N. Shepherd, H. Barr, and N. Stone, "The potential for histological screening using a combination of rapid raman mapping and principal component analysis," *Journal of biophotonics*, vol. 2, no. 1-2, pp. 91–103, 2009.

[14] G. Clemens, J. R. Hands, K. M. Dorling, and M. J. Baker, "Vibrational spectroscopic methods for cytology and cellular research," *Analyst*, vol. 139, no. 18, pp. 4411–4444, 2014.

[15] D. L. Massart, "Chemometrics: a textbook," *Data handling in science and technology*, vol. 2, p. 53, 1988.

[16] C. Krafft, G. Steiner, C. Beleites, and R. Salzer, "Disease recognition by infrared and raman spectroscopy," *Journal of biophotonics*, vol. 2, no. 1-2, pp. 13–28, 2009.

[17] A. Shapiro, O. N. Gofrit, G. Pizov, J. K. Cohen, and J. Maier, "Raman molecular

imaging: a novel spectroscopic technique for diagnosis of bladder cancer in urine specimens," *European urology*, vol. 59, no. 1, pp. 106–112, 2011.

[18] L. T. Kerr, T. M. Lynn, I. M. Cullen, P. J. Daly, N. Shah, S. O'Dea, A. Malkin, and B. M. Hennelly, "Methodologies for bladder cancer detection with raman based urine cytology," *Analytical Methods*, vol. 8, no. 25, pp. 4991–5000, 2016.

[19] L. T. Kerr, K. Domijan, I. Cullen, and B. M. Hennelly, "Applications of raman spectroscopy to the urinary bladder for cancer diagnostics," *Photonics & Lasers in Medicine*, vol. 3, no. 3, pp. 193–224, 2014.

[20] F. Bonnier, D. Traynor, P. Kearney, C. Clarke, P. Knief, C. Martin, J. J. O'Leary, H. J. Byrne, and F. Lyng, "Processing thinprep cervical cytological samples for raman spectroscopic analysis," *Analytical Methods*, vol. 6, no. 19, pp. 7831–7841, 2014.

[21] L. F. C. Carvalho, F. Bonnier, K. O'Callaghan, J. O'Sullivan, S. Flint, H. J. Byrne, and F. M. Lyng, "Raman micro-spectroscopy for rapid screening of oral squamous cell carcinoma," *Experimental and molecular pathology*, vol. 98, no. 3, pp. 502–509, 2015.

[22] A. F. Palonpon, J. Ando, H. Yamakoshi, K. Dodo, M. Sodeoka, S. Kawata, and K. Fujita, "Raman and SERS microscopy for molecular imaging of live cells," *Nature protocols*, vol. 8, no. 4, pp. 677–692, 2013.

[23] L. T. Kerr, K. Domijan, I. Cullen, and B. M. Hennelly, "Applications of raman spectroscopy to the urinary bladder for cancer diagnostics," *Photonics Lasers in Medicine*, vol. 3, no. 3, p. 193–224, 2014.

[24] L. T. Kerr and B. Hennelly, "Methodologies for bladder cancer detection with raman based urine cytology," *Analytical Methods*, vol. 8, no. 25, p. 4991–5000, 2016.

[25] P. Crow, "The use of raman spectroscopy to identify and characterize transitional cell carcinoma in vitro," *BJU international*, vol. 93, no. 9, p. 1232–1236, 2004.

[26] F. M. Lyng, "Vibrational spectroscopy for cervical cancer pathology, from biochemical analysis to diagnostic tool," *Experimental and molecular pathology*, vol. 82, no. 2, p. 121–129, 2007.

[27] L. F. C. Carvalho, "Raman micro-spectroscopy for rapid screening of oral squamous cell carcinoma," *Experimental and molecular pathology*, vol. 98, no. 3, p. 502–509, 2015.

[28] M. Jermyn, "Intraoperative brain cancer detection with raman spectroscopy in humans," *Science translational medicine*, vol. 7, no. 274, p. 274 19–274 19, 2015.

[29] M. S. Bergholt, "Fiber-optic raman spectroscopy probes gastric carcinogenesis in vivo at endoscopy," *Journal of biophotonics*, vol. 6, no. 1, p. 49–59, 2013.

[30] I. P. Santos, E. M. Barroso, T. C. B. Schut, P. J. Caspers, C. G. van Lanschot, D.-H. Choi, M. F. Van Der Kamp, R. W. Smits, R. Van Doorn, R. M. Verdijk, *et al.*, "Raman spectroscopy for cancer detection and cancer surgery guidance: translation to the clinics," *Analyst*, vol. 142, no. 17, pp. 3025–3047, 2017.

[31] L. Lauwerends, H. Abbasi, T. Bakker Schut, P. Van Driel, J. Hardillo, I. Santos, E. Barroso, S. Koljenović, A. Vahrmeijer, R. Baatenburg de Jong, *et al.*, "The complementary value of intraoperative fluorescence imaging and raman spectroscopy for cancer surgery: combining the incompatibles," *European Journal of Nuclear Medicine and Molecular Imaging*, pp. 1–13, 2022.

[32] M. Jermyn, K. Mok, J. Mercier, J. Desroches, J. Pichette, K. Saint-Arnaud, L. Bernstein, M.-C. Guiot, K. Petrecca, and F. Leblond, "Intraoperative brain cancer detection with raman spectroscopy in humans," *Science translational medicine*, vol. 7, no. 274, pp. 274ra19–274ra19, 2015.

[33] I. W. Schie, J. Ruger, A. S. Mondol, A. Ramoji, U. Neugebauer, C. Krafft, and J. Popp, "High-throughput screening raman spectroscopy platform for label-free cellomics," *Analytical chemistry*, vol. 90, no. 3, pp. 2023–2030, 2018.

[34] N. Arend, A. Pittner, A. Ramoji, A. S. Mondol, M. Dahms, J. Rüger, O. Kurzai, I. W. Schie, M. Bauer, J. Popp, *et al.*, "Detection and differentiation of bacterial and fungal infection of neutrophils from peripheral blood using raman spectroscopy," *Analytical Chemistry*, 2020.

[35] A. S. Mondol, S. F. El-Mashtoly, T. Frick, K. Gerwert, J. Popp, and I. W. Schie, "High-content screening raman spectroscopy (hcs-rs) of panitumumab-exposed colorectal cancer cells," *Analyst*, vol. 144, no. 20, pp. 6098–6107, 2019.

[36] K. O'Dwyer, K. Domijan, A. Dignam, M. Butler, and B. M. Hennelly, "Automated raman micro-spectroscopy of epithelial cell nuclei for high-throughput classification," *Cancers*, vol. 13, no. 19, p. 4767, 2021.

[37] K. O'Dwyer, K. Domijan, A. Dignam, Z. Tang, M. Butler, and B. M. Hennelly, "Automated raman cytology for the classification of triple negative breast cancer cell lines," in *Optics in Health Care and Biomedical Optics XI*, vol. 11900, pp. 264–273, SPIE, 2021.

[38] M. G. Shim, "In vivo near-infrared raman spectroscopy: Demonstration of feasibility during clinical gastrointestinal endoscopy," *Photochemistry and photobiology*, vol. 72, no. 1, p. 146–150, 2000.

[39] D. Hutsebaut, P. Vandenabeele, and L. Moens, "Evaluation of an accurate calibration and spectral standardization procedure for raman spectroscopy," *Analyst*, vol. 130, no. 8, pp. 1204–1214, 2005.

[40] J. D. Rodriguez, B. J. Westenberger, L. F. Buhse, and J. F. Kauffman, "Standardization of raman spectra for transfer of spectral libraries across different instruments," *Analyst*, vol. 136, no. 20, pp. 4232–4240, 2011.

[41] T. Dörfer, T. Bocklitz, N. Tarcea, M. Schmitt, and J. Popp, "Checking and improving calibration of raman spectra using chemometric approaches," *Zeitschrift für Physikalische Chemie*, vol. 225, no. 6-7, pp. 753–764, 2011.

[42] T. Bocklitz, T. Dörfer, R. Heinke, M. Schmitt, and J. Popp, "Spectrometer calibration protocol for raman spectra recorded with different excitation wavelengths," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 149, pp. 544–549, 2015.

[43] Y. Chen and L. Dai, "Automated decomposition algorithm for raman spectra based on a voigt line profile model," *Applied optics*, vol. 55, no. 15, pp. 4085–4094, 2016.

[44] S. Guo, C. Beleites, U. Neugebauer, S. Abalde-Cela, N. K. Afseth, F. Alsamad, S. Anand, C. Araujo-Andrade, S. Askrabic, E. Avci, *et al.*, "Comparability of raman spectroscopic configurations: a large scale cross-laboratory study," *Analytical Chemistry*, vol. 92, no. 24, pp. 15745–15756, 2020.

[45] N. Itoh, K. Shirono, and T. Fujimoto, "Baseline assessment for the consistency of raman shifts acquired with 26 different raman systems and necessity of the standardization calibration protocol," *Analytical Sciences*, p. 18P501, 2019.

[46] ASTM E1683-02, "Standard practice for testing the performance of scanning Raman spectrometers," standard, ASTM International, West Conshohocken,Pennsylvania, 2014.

[47] ASTM E2529-06(2014), "Standard guide for testing the resolution of a Raman spectrometer," standard, ASTM International, West Conshohocken,Pennsylvania, 2014.

[48] ASTM E2911-13, "Standard guide for relative intensity correction of Raman spectrometers," standard, ASTM International, West Conshohocken,Pennsylvania, 2013.

[49] A. M. Fales, I. K. Ilev, and T. J. Pfefer, "Evaluation of standardized performance test methods for biomedical raman spectroscopy," *Journal of Biomedical Optics*, vol. 27, no. 7, p. 074705, 2021.

[50] P. McEuen and C. Kittel, "Introduction to solid state physics," 2005.

[51] P. Atkins and T. Overton, *Shriver and Atkins' inorganic chemistry*. Oxford University Press, USA, 2010.

[52] M. Kuball and J. W. Pomeroy, "A review of raman thermography for electronic and opto-electronic device measurement with submicron spatial and nanosecond temporal resolution," *IEEE Transactions on Device and Materials Reliability*, vol. 16, no. 4, pp. 667–684, 2016.

[53] E. Smith and G. Dent, *Modern Raman spectroscopy: a practical approach*. John Wiley & Sons, 2019.

[54] J. R. Ferraro, *Introductory raman spectroscopy*. Elsevier, 2003.

[55] D. Long, "The raman effect,) o (iv, isbn: 978-0-471-49028—9," 2002.

[56] D. Vij and K. Mahesh, *Medical applications of lasers*. Springer Science & Business Media, 2002.

[57] A. N. Zajdel, V. K. Prokofev, S. Raiskii, V. Slavnyi, and E. Y. Schreider, *Tables of spectral lines*. IFI/Plenum Data Corporation, 1970.

[58] T. Jacksier and R. M. Barnes, "Atomic emission spectra of xenon, krypton, and neon: Spectra from 200 to 900 nm by sealed inductively coupled plasma/atomic emission spectroscopy," *Applied spectroscopy*, vol. 48, no. 1, pp. 65–71, 1994.

[59] H.-O. Hamaguchi, "Calibrating multichannel raman spectrometers," *Applied Spectroscopy Reviews*, vol. 24, no. 1-2, pp. 137–174, 1988.

[60] T. Tahara, H. Hamaguchi, and M. Tasumi, "Transient resonance raman spectra of benzophenone and its four isotopic analogues in the lowest excited triplet state," *Journal of Physical Chemistry*, vol. 91, no. 23, pp. 5875–5880, 1987.

[61] K. Iwata, H. Hamaguchi, and M. Tasumi, "Sensitivity calibration of multichannel raman spectrometers using the least-squares-fitted fluorescence spectrum of quinine," *Applied spectroscopy*, vol. 42, no. 1, pp. 12–14, 1988.

[62] C.-H. Tseng, J. F. Ford, C. K. Mann, and T. J. Vickers, "Wavelength calibration of a multichannel spectrometer," *Applied spectroscopy*, vol. 47, no. 11, pp. 1808–1813, 1993.

[63] A. K. Gaigalas, L. Wang, H.-J. He, and P. DeRose, "Procedures for wavelength calibration and spectral response correction of CCD array spectrometers," *Journal of research of the National Institute of Standards and Technology*, vol. 114, no. 4, p. 215, 2009.

[64] F. L. Pedrotti, L. M. Pedrotti, and L. S. Pedrotti, *Introduction to optics*. Cambridge University Press, 2017.

[65] J. Lerner and A. Thevenon, "The optics of spectroscopy," *Tutorial Jobin-Yvon*, 1988.

[66] E. Palmer, M. Hutley, A. Franks, J. Verrill, and B. Gale, "Diffraction gratings (manufacture)," *Reports on Progress in Physics*, vol. 38, no. 8, p. 975, 1975.

[67] S. P. Davis and S. P. Davis, *Diffraction grating spectrographs*. Holt, Rinehart, and Winston New York, 1970.

[68] T. J. Vickers, C. A. Rosen, and C. K. Mann, "Compact raman spectrometers: data handling methods," *Applied spectroscopy*, vol. 50, no. 8, pp. 1074–1081, 1996.

[69] D. A. Carter, W. R. Thompson, C. E. Taylor, and J. E. Pemberton, "Frequency/wavelength calibration of multipurpose multichannel raman spectrometers. part ii: calibration fit considerations and calibration standards," *Applied spectroscopy*, vol. 49, no. 11, pp. 1561–1576, 1995.

[70] P. Martinsen, B. Jordan, A. McGlone, P. Gaastra, and T. Laurie, "Accurate and precise wavelength calibration for wide bandwidth array spectrometers," *Applied spectroscopy*, vol. 62, no. 9, pp. 1008–1012, 2008.

[71] H.-K. Fu, Y.-L. Liu, T.-T. Chen, C.-P. Wang, and P.-T. Chou, "The study of spectral correction algorithm of charge-coupled device array spectrometer," *IEEE Transactions on Electron Devices*, vol. 61, no. 11, pp. 3796–3802, 2014.

[72] Y. Sun, C. Huang, G. Xia, S. Jin, and H. Lu, "Accurate wavelength calibration method for compact ccd spectrometer," *JOSA A*, vol. 34, no. 4, pp. 498–505, 2017.

[73] S. Nevas, G. Wübbeler, A. Sperling, C. Elster, and A. Teuber, "Simultaneous correction of bandpass and stray-light effects in array spectroradiometer data," *Metrologia*, vol. 49, no. 2, p. S43, 2012.

[74] J. Dubrovkin, "Identification of peak positions using second-order derivative spectra and tikhonov deconvolution method. comparison study," *Journal of Emerging Technologies in Computational and Applied Sciences*, vol. 2, p. 10, 2014.

[75] E. R. Woolliams, R. Baribeau, A. Bialek, and M. G. Cox, "Spectrometer bandwidth correction for generalized bandpass functions," *Metrologia*, vol. 48, no. 3, p. 164, 2011.

[76] M. Morháč and V. Matoušek, "Complete positive deconvolution of spectrometric data," *Digital Signal Processing*, vol. 19, no. 3, pp. 372–392, 2009.

[77] T. Sundius, "Computer fitting of voigt profiles to raman lines," *Journal of Raman Spectroscopy*, vol. 1, no. 5, pp. 471–488, 1973.

[78] J. A. Holy, "Determination of spectrometer-detector parameters from calibration spectra and the use of the parameters in spectrometer calibrations," *Applied spectroscopy*, vol. 58, no. 10, pp. 1219–1227, 2004.

[79] K. Liu and F. Yu, "Accurate wavelength calibration method using system parameters for grating spectrometers," *Optical Engineering*, vol. 52, no. 1, p. 013603, 2013.

[80] F. Zhang, C. Chen, J. Liu, and Z. Wang, "Error analysis of mechanical system

and wavelength calibration of monochromator," *Review of Scientific Instruments*, vol. 89, no. 2, p. 023112, 2018.

[81] F. Scotti and R. E. Bell, "High accuracy wavelength calibration for a scanning visible spectrometer," *Review of Scientific Instruments*, vol. 81, no. 10, p. 10D732, 2010.

[82] R. E. Bell and F. Scotti, "High-throughput accurate-wavelength lens-based visible spectrometer," *Review of Scientific Instruments*, vol. 81, no. 10, p. 10D731, 2010.

[83] K. Birch and M. Downs, "An updated edlén equation for the refractive index of air," *Metrologia*, vol. 30, no. 3, p. 155, 1993.

[84] L. Yuan and L. Qiu, "Wavelength calibration methods in laser wavelength measurement," *Applied Optics*, vol. 60, no. 15, pp. 4315–4324, 2021.

[85] P. E. Ciddor, "Refractive index of air: new equations for the visible and near infrared," *Applied optics*, vol. 35, no. 9, pp. 1566–1573, 1996.

[86] D. A. Sadler, D. Littlejohn, and C. V. Perkins, "Automatic wavelength calibration procedure for use with an optical spectrometer and array detector," *Journal of Analytical Atomic Spectrometry*, vol. 10, no. 3, pp. 253–257, 1995.

[87] D. L. Miller and A. Scheeline, "A computer program for the collection, reduction and analysis of echelle spectra," *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 48, no. 8, pp. E1053–E1062, 1993.

[88] A. Scheeline, C. A. Bye, D. L. Miller, S. W. Rynders, and R. C. Owen Jr, "Design and characterization of an echelle spectrometer for fundamental and applied emission spectrochemical analysis," *Applied spectroscopy*, vol. 45, no. 3, pp. 334–346, 1991.

[89] B. Hennelly and D. Liu, "Improved wavelength calibration by modelling the spectrometer," *Applied Spectroscopy*, p. 00037028221111796, 2022.

[90] T. J. Vickers and C. K. Mann, "Raman shift calibration of a compact multichannel spectrometer," *Applied Spectroscopy*, vol. 53, no. 12, pp. 1617–1622, 1999.

[91] R. Hawkins, M. Hoke, and J. Shaw, "Wavenumber calibration of fourier transform spectra," *Applied Spectroscopy*, vol. 37, no. 2, pp. 134–139, 1983.

[92] T. J. Vickers and C. K. Mann, "On-line monitoring by raman spectroscopy: instrument control and calibration," in *Optical Sensors for Environmental and Chemical Process Monitoring*, vol. 2367, pp. 219–227, International Society for Optics and Photonics, 1995.

[93] A. W. Fountain III, C. K. Mann, and T. J. Vickers, "Routine wavenumber calibration of an ft-raman spectrometer," *Applied spectroscopy*, vol. 49, no. 7, pp. 1048–1053, 1995.

[94] R. W. Berg and T. Nørbygaard, "Wavenumber calibration of ccd detector raman spectrometers controlled by a sinus arm drive," *Applied Spectroscopy Reviews*, vol. 41, no. 2, pp. 165–183, 2006.

[95] D. V. Petrov, I. Matrosov, D. Sedinkin, and A. Zaripov, "Wavenumber calibration of a multichannel raman spectrometer," in *23rd International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics*, vol. 10466, p. 1046606, International Society for Optics and Photonics, 2017.

[96] A. W. Fountain, T. J. Vickers, and C. K. Mann, "Factors that affect the accuracy of raman shift measurements on multichannel spectrometers," *Applied spectroscopy*, vol. 52, no. 3, pp. 462–468, 1998.

[97] A. Raj, C. Kato, H. A. Witek, and H.-o. Hamaguchi, "Toward standardization of raman spectroscopy: Accurate wavenumber and intensity calibration using rotational raman spectra of h2, hd, d2, and vibration–rotation spectrum of o2," *Journal of Raman Spectroscopy*, vol. 51, no. 10, pp. 2066–2082, 2020.

[98] S. Wollman and P. Bohn, "Evaluation of polynomial fitting functions for use with

ccd arrays in raman spectroscopy," *Applied Spectroscopy*, vol. 47, no. 1, pp. 125–126, 1993.

[99] O. Ryabchykov, S. Guo, and T. Bocklitz, "Analyzing raman spectroscopic data," *Physical Sciences Reviews*, vol. 4, no. 2, 2019.

[100] J. C. Weatherall, J. Barber, C. S. Brauer, T. J. Johnson, Y.-F. Su, C. D. Ball, B. T. Smith, R. Cox, R. Steinke, P. McDaniel, *et al.*, "Adapting raman spectra from laboratory spectrometers to portable detection libraries," *Applied Spectroscopy*, vol. 67, no. 2, pp. 149–157, 2013.

[101] M. M. Carrabba, "Wavenumber standards for r aman spectrometry," *Handbook of vibrational spectroscopy*, 2006.

[102] O. E. COUNCIL, "European pharmacopoeia.," *Strasbourg, France: Council of Europe*, 2016.

[103] C. Shen, T. J. Vickers, and C. K. Mann, "Abscissa error detection and correction in raman spectroscopy," *Applied spectroscopy*, vol. 46, no. 5, pp. 772–777, 1992.

[104] M. Pelletier, "Effects of temperature on cyclohexane raman bands," *Applied spectroscopy*, vol. 53, no. 9, pp. 1087–1096, 1999.

[105] A. E1840-96, "Standard guide for raman shift standards for spectrometer calibration," 2014.

[106] J. Komasa, K. Piszczatowski, G. Łach, M. Przybytek, B. Jeziorski, and K. Pachucki, "Quantum electrodynamics effects in rovibrational spectra of molecular hydrogen," *Journal of chemical theory and computation*, vol. 7, no. 10, pp. 3105–3115, 2011.

[107] K. Pachucki and J. Komasa, "Rovibrational levels of hd," *Physical Chemistry Chemical Physics*, vol. 12, no. 32, pp. 9188–9196, 2010.

[108] H. Edwards, D. Long, K. Najm, and M. Thomsen, "The vibration-rotation raman spectra of 18o2, 17o18o, 17o2 and 16o2," *Journal of Raman Spectroscopy*, vol. 10, no. 1, pp. 60–63, 1981.

[109] M. M. Carrabba, "Wavenumber standards for raman spectrometry," in *Handbook of vibrational spectroscopy*, 2006.

[110] R. Jones and A. Nadeau, "Further observations on the use of indene for the wavenumber calibration of infrared spectrometers," *Spectrochimica Acta*, vol. 20, no. 7, p. 1175–1183, 1964.

[111] R. W. Berg and T. Nørbygaard, "Wavenumber calibration of ccd detector raman spectrometers controlled by a sinus arm drive," *Applied Spectroscopy Reviews*, vol. 41, no. 2, p. 165–183, 2006.

[112] A. W. Fountain, III, C. K. Mann, and T. J. Vickers, "Routine wavenumber calibration of an ft-raman spectrometer," *Applied spectroscopy*, vol. 49, no. 7, p. 1048–1053, 1995.

[113] J.-G. Skinner and W. Nilsen, "Absolute raman scattering cross-section measurement of the 992 cm 1 line of benzene," *JOSA*, vol. 58, no. 1, p. 113–119, 1968.

[114] P. Hendra and E.-J. Loader, "Laser raman spectra of adsorbed species," *Transactions of the Faraday Society*, vol. 67, p. 828–840, 1971.

[115] D. Boor and Carl, *A practical guide to splines*, vol. 27. New York: Springer-Verlag, 1978.

[116] H. Chen, "Automatic standardization method for raman spectrometers with applications to pharmaceuticals," *Journal of Raman Spectroscopy*, vol. 46, no. 1, p. 147–154, 2015.

[117] D. Dussault and P. Hoess, "Noise performance comparison of iccd with ccd and emccd cameras," *Infrared Systems and Photoelectronic Technology*, vol. 5563, 2004.

[118] M. T. Murphy, T. Udem, R. Holzwarth, A. Sizmann, L. Pasquini, C. Araujo-Hauck, H. Dekker, S. D'Odorico, M. Fischer, T. Hänsch, *et al.*, "High-precision wavelength calibration of astronomical spectrographs with laser frequency combs," *Monthly Notices of the Royal Astronomical Society*, vol. 380, no. 2, pp. 839–847, 2007.

[119] G. Wetzstein, I. Ihrke, D. Lanman, and W. Heidrich, "Computational plenoptic imaging," in *Computer Graphics Forum*, vol. 30, pp. 2397–2426, Wiley Online Library, 2011.

[120] C. Ma, H. Lin, G. Zhang, and R. Du, "An efficient calibration method for multi-spectral imaging," *Optics Communications*, vol. 420, pp. 14–25, 2018.

[121] M. Szkulmowski, S. Tamborski, and M. Wojtkowski, "Spectrometer calibration for spectroscopic fourier domain optical coherence tomography," *Biomedical optics express*, vol. 7, no. 12, pp. 5042–5054, 2016.

[122] A. Bradu, S. Rivet, and A. Podoleanu, "Master/slave interferometry–ideal tool for coherence revival swept source optical coherence tomography," *Biomedical Optics Express*, vol. 7, no. 7, pp. 2453–2468, 2016.

[123] F. Lange and I. Tachtsidis, "Clinical brain monitoring with time domain nirs: a review and future perspectives," *Applied Sciences*, vol. 9, no. 8, p. 1612, 2019.

[124] T. Scheeren, P. Schober, and L. Schwarte, "Monitoring tissue oxygenation by near infrared spectroscopy (nirs): background and current applications," *Journal of clinical monitoring and computing*, vol. 26, no. 4, pp. 279–287, 2012.

[125] D. Alomar, C. Gallo, M. Castaneda, and R. Fuchslocher, "Chemical and discriminant analysis of bovine meat by near infrared reflectance spectroscopy (nirs)," *Meat science*, vol. 63, no. 4, pp. 441–450, 2003.

[126] J. S. Shenk and M. O. Westerhaus, "New standardization and calibration procedures for nirs analytical systems," *Crop Science*, vol. 31, no. 6, pp. 1694–1696, 1991.

[127] K. W. Busch, O. Soyemi, D. Rabbe, K. Humphrey, B. Dundee, and M. A. Busch, "Wavelength calibration of a dispersive near-infrared spectrometer using trichloromethane as a calibration standard," *Applied Spectroscopy*, vol. 54, no. 9, pp. 1321–1326, 2000.

[128] Í. P. Caliari, J. V. Nicacio, M. H. P. Barbosa, and R. F. Teófilo, "Reconsidering the need for empirical alignment and wavelength calibration steps in the building of a dispersive nir spectrometer with an application for ethanol quantification using a polymer filament 3d printer," *Analytical Chemistry*, vol. 93, no. 33, pp. 11388–11397, 2021.

[129] T. Vankeirsbilck, A. Vercauteren, W. Baeyens, G. Van der Weken, F. Verpoort, G. Vergote, and J. P. Remon, "Applications of raman spectroscopy in pharmaceutical analysis," *TrAC trends in analytical chemistry*, vol. 21, no. 12, pp. 869–877, 2002.

[130] A. C. McGovern, D. Broadhurst, J. Taylor, N. Kaderbhai, M. K. Winson, D. A. Small, J. J. Rowland, D. B. Kell, and R. Goodacre, "Monitoring of complex industrial bioprocesses for metabolite concentrations using modern spectroscopies and machine learning: application to gibberellic acid production," *Biotechnology and Bioengineering*, vol. 78, no. 5, pp. 527–538, 2002.

[131] K. Kneipp, H. Kneipp, I. Itzkan, R. R. Dasari, and M. S. Feld, "Ultrasensitive chemical analysis by raman spectroscopy," *Chemical reviews*, vol. 99, no. 10, pp. 2957–2976, 1999.

[132] P. Lasch, "Spectral pre-processing for biomedical vibrational spectroscopy and microspectroscopic imaging," *Chemometrics and Intelligent Laboratory Systems*, vol. 117, pp. 100–114, 2012.

[133] W. Neumann, "Fundamentals of dispersive optical spectroscopy systems," Society of Photo-Optical Instrumentation Engineers, 2013.

[134] J. Zhao, "Image curvature correction and cosmic removal for high-throughput

dispersive raman spectroscopy," *Applied spectroscopy*, vol. 57, no. 11, pp. 1368–1375, 2003.

[135] C. D. Allemand, "Coma correction in czerny–turner spectrographs," *JOSA*, vol. 58, no. 2, pp. 159–163, 1968.

[136] P. Gemperline, *Practical guide to chemometrics*. CRC press, 2006.

[137] R. Wehrens, *Chemometrics with R*, vol. 3. Springer, 2011.

[138] J. E. Harvey and C. L. Vernold, "Description of diffraction grating behavior in direction cosine space," *Applied optics*, vol. 37, no. 34, pp. 8158–8159, 1998.

[139] J. Wells, "lorentzfit," https://mathworks.com/matlabcentral/fileexchange/33775-lorentzfit-x-y-varargin, journal=MATLAB central file exchange, year=2020.

[140] R. Alford, H. M. Simpson, J. Duberman, G. C. Hill, M. Ogawa, C. Regino, H. Kobayashi, and P. L. Choyke, "Toxicity of organic fluorophores used in molecular imaging: literature review," *Molecular imaging*, vol. 8, no. 6, pp. 7290–2009, 2009.

[141] M. Ojanen, P. Kärhä, and E. Ikonen, "Spectral irradiance model for tungsten halogen lamps in 340–850 nm wavelength range," *Applied optics*, vol. 49, no. 5, pp. 880–886, 2010.

[142] B. Hennelly and D. Liu, "Wavenumber calibration by modelling the spectrometer," *submitted to Journal of Raman Spectroscopy*, 2022.

[143] L. T. Kerr and B. Hennelly, "A multivariate statistical investigation of background subtraction algorithms for raman spectra of cytology samples recorded on glass slides," *Chemometrics and Intelligent Laboratory Systems*, vol. 158, pp. 61–68, 2016.

[144] L. T. Kerr, H. J. Byrne, and B. M. Hennelly, "Optimal choice of sample substrate and laser wavelength for raman spectroscopic analysis of biological specimen," *Analytical Methods*, vol. 7, no. 12, pp. 5041–5052, 2015.

[145] H. Takeuchi, S. Hashimoto, and I. Harada, "Simple and efficient method to eliminate spike noise from spectra recorded on charge-coupled device detectors," *Applied spectroscopy*, vol. 47, no. 1, pp. 129–131, 1993.

[146] E. Galvez and P. Koch, "Use of four mirrors to rotate linear polarization but preserve input–output collinearity. ii.," *JOSA A*, vol. 14, no. 12, pp. 3410–3414, 1997.

[147] A. Zaidel', V. Prokof'ev, S. Raiskii, V. Slavnyi, and E. Y. Shreider, *Tables of Spectral Lines*. Springer, 1970.

[148] J. Dong, M. Hong, Y. Xu, and X. Zheng, "A practical convolutional neural network model for discriminating raman spectra of human and animal blood," *Journal of Chemometrics*, vol. 33, no. 11, p. e3184, 2019.

[149] T. J. Vickers, R. E. Wambles Jr, and C. K. Mann, "Curve fitting and linearity: data processing in raman spectroscopy," *Applied Spectroscopy*, vol. 55, no. 4, pp. 389–393, 2001.

[150] B. D. Beier and A. J. Berger, "Method for automated background subtraction from raman spectra containing known contaminants," *Analyst*, vol. 134, no. 6, pp. 1198–1202, 2009.

[151] F. Xinwei, Z. Zhongliang, S. Mengjie, and C. Peisheng, "The method of baseline drift correction of raman spectrum based on polynomial fitting," *Computers and Applied Chemistry*, vol. 6, 2009.

[152] J. Zhao, H. Lui, D. I. McLean, and H. Zeng, "Automated autofluorescence background subtraction algorithm for biomedical raman spectroscopy," *Applied spectroscopy*, vol. 61, no. 11, pp. 1225–1232, 2007.

[153] C. A. Lieber and A. Mahadevan-Jansen, "Automated method for subtraction of fluorescence from biological raman spectra," *Applied spectroscopy*, vol. 57, no. 11, pp. 1363–1367, 2003.

[154] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[155] J. Liu, M. Osadchy, L. Ashton, M. Foster, C. J. Solomon, and S. J. Gibson, "Deep convolutional neural networks for raman spectrum recognition: a unified solution," *Analyst*, vol. 142, no. 21, pp. 4067–4074, 2017.

[156] J. Acquarelli, T. van Laarhoven, J. Gerretzen, T. N. Tran, L. M. Buydens, and E. Marchiori, "Convolutional neural networks for vibrational spectroscopic data analysis," *Analytica chimica acta*, vol. 954, pp. 22–31, 2017.

[157] C. Cui and T. Fearn, "Modern practical convolutional neural networks for multivariate regression: Applications to nir calibration," *Chemometrics and Intelligent Laboratory Systems*, vol. 182, pp. 9–20, 2018.

[158] S. Malek, F. Melgani, and Y. Bazi, "One-dimensional convolutional neural networks for spectroscopic signal regression," *Journal of Chemometrics*, vol. 32, no. 5, p. e2977, 2018.

[159] S. L. Neal, "Multivariate analysis of mixed lipid aggregate phase transitions monitored using raman spectroscopy," *Applied spectroscopy*, vol. 72, no. 1, pp. 102–113, 2018.

[160] W. Lee, A. T. Lenferink, C. Otto, and H. L. Offerhaus, "Classifying raman spectra of extracellular vesicles based on convolutional neural networks for prostate cancer detection," *Journal of Raman spectroscopy*, vol. 51, no. 2, pp. 293–300, 2020.

[161] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," *arXiv preprint arXiv:1506.00019*, 2015.

[162] Y. Bengio, P. Simard, and P. Frasconi, “Learning long-term dependencies with gradient descent is difficult,” *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 157–166, 1994.

[163] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[164] C. Olah, “Understanding LSTM networks, 2015,” *URL http://colah. github. io/posts/2015-08-Understanding-LSTMs*, vol. 19, pp. 1–19, 2015.

[165] S. Trasatti, “"Section 13.3 line widths", atkins’ physical chemistry, p. atkins, j. de paula, oxford university press, oxford, uk (2006), isbn: 0198700725,” 2007.

[166] C. P. Poole Jr, *"Line Shape", Encyclopedic dictionary of condensed matter physics*. Academic Press, 2004.