Titre: Title:	Bringing Real-Time Raman Spectroscopy Tissue Characterisation to in Vivo Applications Including Cancer Diagnosis
Auteur: Author:	Guillaume Sheehy
Date:	2025
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Sheehy, G. (2025). Bringing Real-Time Raman Spectroscopy Tissue Characterisation to in Vivo Applications Including Cancer Diagnosis [Thèse de doctorat, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/65813/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/65813/
Directeurs de recherche: Advisors:	Frédéric Leblond
Programme: Program:	Génie biomédical

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Bringing Real-Time Raman Spectroscopy Tissue Characterisation to In Vivo Applications Including Cancer Diagnosis

GUILLAUME SHEEHY
Institut de génie biomédical

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*

Génie biomédical

Mai 2025

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Cette thèse intitulée :

Bringing Real-Time Raman Spectroscopy Tissue Characterisation to In Vivo Applications Including Cancer Diagnosis

présentée par **Guillaume SHEEHY**
en vue de l'obtention du diplôme de *Philosophiæ Doctor*
a été dûment acceptée par le jury d'examen constitué de :

Jean PROVOST, président
Frédéric LEBLOND, membre et directeur de recherche
Caroline BOUDOUX, membre
James TUNNELL, membre externe

DEDICATION

We are standing on the shoulders of giants,

mine are named Patricia and Sylvain.

To Nicolas

To Catherine

ACKNOWLEDGEMENTS

First and foremost, I wish to express my deepest gratitude to my research director, Professor Frédéric Leblond. Words are scarcely sufficient to articulate the profound privilege it has been to conduct my doctoral research under the guidance of such an inspiring leader, dedicated mentor, and exceptional scientist. Over a decade ago, Frédéric granted me the opportunity to undertake a research internship. That opportunity blossomed into a long-standing collaboration, culminating in both a Master's degree and now this doctoral work. His impact as a lab director, and his capacity to lead a multidisciplinary research team with exceptional vision and rigour, are a testament not only to his world-class scientific acumen but also to his unwavering passion and deep commitment to the personal and professional development of his students. His investment in our collective success resonates profoundly and has been transformative for all who have had the privilege of working under his direction.

I also wish to acknowledge the past and present members of the LUMED Laboratory. Each of you, in your own way, has contributed to the scientific and personal growth I have experienced during this journey. I extend a heartfelt thank you to Fabien, a lifelong friend and colleague whose presence and shared experiences remain deeply cherished. To Fredo, whose long-standing contributions across multiple projects have left an indelible mark on the lab's evolution—thank you.

I also wish to express my appreciation to Trang and Katie for their camaraderie and collaboration, and to Audrey, François, and Sandryne. Further thanks go to Esmat, David, Samaneh, Hugo, Alice, Victor, Antoine, Gabriel, Laurence, Tommy, Elias, Najib, Billie, Jean-Philippe, and Émile, whose engagement and dedication to our collective efforts have enriched this work. A sincere thank you to the many interns over the years, whose enthusiasm and curiosity have been both motivating and inspiring.

To Simon, Dominique, Estelle, and Amélie—your steadfast friendship and encouragement have been invaluable.

To my spouse, Catherine—this dissertation is as much yours as it is mine. The pursuit of a PhD demands significant sacrifices, not only from the individual but from those closest to them. Your understanding, patience, and boundless support have sustained me through every challenge. Your belief in me has been a source of strength, and I owe a profound debt of gratitude to you for being my partner, my anchor, and my unwavering support.

To my brother Nicolas, thank you for the countless stimulating conversations on education,

academia, and technology. Your insight, perspective and experience have informed my thinking in ways that are deeply woven into the fabric of this work.

To my parents, Patricia and Sylvain—no expression of thanks could ever adequately capture the impact you have had on my life. You instilled in me the values of perseverance, selflessness, and intellectual curiosity—values that have guided me unfailingly through the complexities of this academic endeavour. When push comes to shove and when difficulty arises, I will remind myself that when a goal is set, it is "non négociable" and that one does not quit.

To my grandfather, whose memory I carry with great affection and gratitude. Though he did not live to see the completion of this work, his influence is present in every circuit diagram and every soldered joint. He was the one who first sparked my interest in electronics, physics, and the inner workings of technology. I fondly recall the countless hours spent exploring the electronic components and PCBs he brought home from work, the thrill of receiving my first electrical shock under his amused supervision, and the patience with which he taught me to solder. The practical knowledge he imparted and the curiosity he nurtured in me have profoundly shaped my professional identity. I know he would have been immensely proud, as he always was.

Finally, I wish to acknowledge the institutional and financial support without which this research would not have been possible. I am grateful to Polytechnique Montréal, the Centre de Recherche du Centre Hospitalier de l'Université de Montréal, NSERC, TransMedTech, FRQNT, CIHR, and Mitacs. Their contributions, both direct and indirect, have been foundational to the completion of this doctoral work.

To everyone else that I forgot to explicitly mentioned, but should have — Thank you!!!

RÉSUMÉ

La spectroscopie Raman est une technique d'analyse optique sans agent de contraste qui permet d'identifier la composition moléculaire des tissus biologiques. Malgré un fort potentiel pour des applications diagnostiques en temps réel, son intégration en milieu clinique demeure limitée. Cela s'explique notamment par la complexité des systèmes existants, l'absence d'outils de traitement standardisés accessibles et la dépendance à des solutions logicielles et matérielles propriétaires. Cette thèse propose une plateforme complète, modulaire et open source, spécifiquement développée pour les besoins de la recherche biomédicale appliquée et des milieux cliniques.

Le projet se structure autour de deux axes principaux. D'abord, la bibliothèque logicielle *Open Raman Processing Library* (ORPL) a été conçue pour fournir un ensemble d'outils ouverts et robustes de traitement de spectres Raman, y compris un nouvel algorithme de soustraction de fluorescence, *BubbleFill*. Cet algorithme s'est montré particulièrement efficace pour les signaux caractérisés par un faible rapport signal-sur-bruit, commun en *in vivo* pour les tissus biologiques. ORPL est distribuée sous licence MIT et est maintenant utilisée comme outil de traitement standard dans les projets de spectroscopie Raman du groupe de recherche Lumed.

Ensuite, le logiciel *Open Raman Acquisition Software* (ORAS) a été développé afin de permettre l'acquisition de données Raman dans un cadre clinique. ORAS regroupe des modules de contrôle automatisé (exposition, accumulation), une visualisation en temps réel des spectres, des profils d'acquisition prédéfinis conformes aux normes de sécurité laser, ainsi qu'un module d'inférence pour la classification automatique des tissus. L'ensemble est conçu pour être utilisé efficacement par du personnel non technique, tout en assurant une traçabilité complète des acquisitions.

La plateforme a été concrètement implantée au travers du *Lumed Raman System*, composé d'instruments commerciaux modulaires et est complétée par des composantes électroniques, des pièces fabriquées en impression 3D et un micro-contrôleur conçu sur mesure pour assurer la sécurité laser. Quatre systèmes ont été fabriqués et déployés dans des établissements hospitaliers à Montréal. Ces systèmes sont actuellement utilisés dans des études cliniques impliquant la collecte de données *in vivo* en salle d'opération, notamment lors de chirurgies du cerveau chez l'adulte et l'enfant. Les performances ont été comparées à celles d'un système près de la commercialisation (ODS/Reveal Sentry), démontrant des bénéfices tangibles en termes de qualité de signal, de robustesse opérationnelle et d'ergonomie logicielle.

Tout le matériel logiciel, les plans techniques, les guides de fabrication et la documentation

sont accessibles publiquement suivant l'esprit open source. En plus des aspects techniques, la thèse discute des limites actuelles du système — par exemple, la difficulté d'automatiser le choix des paramètres de traitement des spectres ou la dépendance à certaines composantes propriétaires — et propose des pistes concrètes pour les travaux à venir, notamment l'intégration du traitement dans les chaines d'apprentissage automatique, le développement de matériaux de calibration stérilisables, et la création de ressources de formation approchables et accessibles.

En somme, cette thèse contribue à rapprocher la spectroscopie Raman d'un usage clinique concret. Elle offre un exemple de plateforme construite dès le départ suivant des principes d'ouverture, de modularité et de transférabilité. Elle fournit également une base solide pour le développement de futures solutions optiques de diagnostic dans un cadre médical canadien et international.

ABSTRACT

Raman spectroscopy offers a powerful, label-free method for probing the molecular composition of biological tissues. Despite its demonstrated diagnostic potential, widespread adoption of the technology in clinical environments has been hindered by several persistent limitations, including high system complexity, lack of standardized data processing workflows, limited integration with clinical procedures, and the proprietary nature of most commercial solutions. This thesis addresses these challenges through the development of a comprehensive, modular, and open-source Raman spectroscopy platform specifically designed for clinical research and intraoperative applications.

The work is organized around two central contributions. First, the Open Raman Processing Library (ORPL) is introduced as a robust and extensible Python-based software framework for preprocessing Raman spectra acquired from CCD-based dispersive spectrometers. The library includes modules for dark signal subtraction, cosmic ray removal, wavenumber calibration, baseline correction, normalization, and smoothing. Notably, a novel baseline correction algorithm—BubbleFill—is developed and evaluated. This method leverages iterative pseudo-convex envelope (e.g. bubbles) fitting to remove fluorescence backgrounds and is shown to outperform state-of-the-art alternatives, particularly in spectra with low Raman-to-fluorescence signal ratios, as typically observed in *in-vivo* biological measurements. The ORPL library is released under the permissive MIT license and is now the standard processing framework for all Raman spectroscopy projects within the Lumed research group.

The second major contribution is the Open Raman Acquisition Software (ORAS), a fully integrated, user-oriented acquisition platform designed to control all aspects of Raman data collection in clinical settings. ORAS provides real-time spectrum visualization, automated acquisition routines—including Automatic Exposure Control (AEC) and Automatic Number of Accumulations Optimization (ANO)—and supports predefined acquisition profiles to ensure compliance with laser safety regulations. In addition, ORAS features real-time classification capabilities through a modular AI inference interface that loads externally defined Python-based models and presents diagnostic predictions to clinical users in a clear and interpretable manner. The integration of ORPL within ORAS enables seamless transition from raw signal acquisition to processed, classified output within a single software environment.

The complete platform was realized in hardware as the Lumed Raman System, a modular Raman device constructed from commercially available components, supplemented by custom-designed enclosures, electronics, and safety systems. Four Lumed systems were built

and deployed in clinical environments across Montreal. These systems have been used in multiple ongoing clinical studies, with *in-vivo* measurements acquired during neurosurgical procedures on adult and pediatric brain tissue. Comparative analyses against a leading pre-commercial Raman system (ODS/Reveal Sentry) demonstrate improved spectral quality and usability of the Lumed platform, with superior dynamic range utilization and no saturation events, attributed to the automated control features of ORAS.

Additional contributions of this thesis include the development of standardized communication protocols for Raman system components, integration of metadata capture within the acquisition interface, and the release of comprehensive documentation and source code to support reproducibility and community adoption. Limitations encountered during this work—such as the challenge of automating spectral processing parameter selection and dependence on proprietary hardware drivers—are discussed, along with concrete recommendations for future development. These include the use of machine learning pipelines to jointly optimize preprocessing and classification, the design of sterilizable calibration standards, and the prioritization of educational resources to support user training and system deployment.

Taken together, the developments presented in this thesis contribute to the maturation of Raman spectroscopy as a viable clinical tool. By adopting an open and modular approach at every level—from signal processing to hardware design—this work supports a growing ecosystem of accessible and extensible tools for biomedical optics. It also serves as a practical demonstration of how translational technologies can be built to balance clinical rigor, engineering robustness, and research openness. As the field moves toward more scalable and clinically integrated solutions, the platform developed here is well-positioned to serve as both a foundation and a catalyst for future innovations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AEC	Automatic Exposure Control
AI	Artificial Intelligence
ANO	Automatic N Optimization
ASSI	Average Signed Squared Intensity
BCS	Breast Conserving Surgery
CARS	Coherent Anti-Stokes Raman Scattering
CEUS	Contrast-Enhanced Ultrasound
CLE	Confocal Laser Endomicroscopy
EHR	Electronic Health Records
FITC	Fluorescein Isothiocyanate
FLIM	Fluorescence Lifetime Imaging
GPL	GNU General Public License
HTA	Health Technology Assessment
ICG	Indocyanine Green
LRS	Lumed Raman System
MIT	Massachusetts Institute of Technology (license)
ML	Machine Learning
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
nMSE	Normalized Mean Squared Error
NIST	National Institute of Standards and Technology
OCT	Optical Coherence Tomography
ORAS	Open Raman Acquisition Software
ORPL	Open Raman Processing Library
OSS	Open-Source Software
PDMS	Polydimethylsiloxane
QF	Quality Factor
RBR	Raman to Baseline Ratio
RS	Raman Spectroscopy
SBR	Signal to Background Ratio
SCPI	Standard Commands for Programmable Instruments
SDK	Software Development Kit
SERS	Surface-Enhanced Raman Spectroscopy

SERDS	Shifted Excitation Raman Difference Spectroscopy
SNV	Standard Normal Variate
SNR	Signal to Noise Ratio
SRM	Standard Reference Material
SVM	Support Vector Machine
TRFS	Time-Resolved Fluorescence Spectroscopy

LIST OF APPENDIXES

CHAPTER 1 INTRODUCTION

1.1 Motivation

Raman spectroscopy (RS) has emerged as a powerful, non-invasive and label-free optical technique usable for medical diagnosis. It features a high biochemical specificity while requiring minimal sample preparation [7, 8]. It can be effective in multiple context such as the detection and monitoring of various diseases, including cancer [9, 10, 11] and neurodegenerative pathology [12, 13]. Furthermore, it was shown to provide detailed molecular information in the analysis of other biological tissues, including bacteria[14] and biofluids[15, 7, 16, 17, 3]. Researchers in the field have combined RS with multivariate and statistical analysis for effective pattern recognition and sample classification [18]. This, along with its potential for in-vivo and real-time measurements using custom fibre-optic probes, make it particularly well suited for clinical applications [7]. Furthermore, the dramatic events following the recent world pandemic of Sars CoV-2 have renewed interest for scalable, fast and reagent-free methods for the detection of infectious diseases. Since then, RS was identified as a potential screening tool for COVID-19 and other infectious diseases [3, 12, 13].

All in all, RS technologies appear to possess everything needed to succeed as a powerful modern diagnostic tool adaptable to many facets of health care. However, promising as they may be, commercially available Raman diagnostic systems have so far failed to translate this potential into widespread clinical use. Several key challenges remain, most notably the high cost of instrumentation and maintenance, the requirement for advanced training of acquisition system operators and the need for sophisticated data analysis expertise. The latter may constitute one of the harder barriers to overcome for RS to reach broader acceptance among clinicians and cancer scientists [19]. Furthermore, the relatively low signal strength inherent in RS measurements, particularly when analyzing complex biological samples, necessitates highly sensitive detectors and optimized experimental protocols [20]. These factors further contribute to the technical difficulties and variability of signal quality observable throughout the published literature. The development of robust, miniaturized probes suitable for point-of-care applications is also still ongoing, often requiring a custom design for each application. Despite these hurdles, continued advancements in laser technology, detector sensitivity, and data processing algorithms are steadily improving the capabilities of RS. With further refinement, RS could eventually overcome these limitations and become a truly integrated component of modern medical diagnostics.

The creation of open-source platforms could fundamentally shift and revitalize the landscape

of RS research and spearhead a new wave of progress toward clinical integration. Driven by community-wide contributions, such platforms would foster collaborative development, dramatically reduce redundant efforts across academic research groups and accelerate innovation. Specifically, open-source software development allows for rapid iteration and customization tailored to diverse research needs, democratizing access to sophisticated processing techniques previously limited to groups already established in the field. This shift could lead to significant standardization of data acquisition protocols and analysis methods, mitigating the variability observed in existing literature and ultimately improving the reliability of RS results. Furthermore, shared open-source platforms would naturally facilitate the construction of comprehensive, standardized databases - accessible to all researchers - representing a crucial step towards translating RS's potential into truly impactful clinical diagnostics.

1.2 Research Objectives

In light of the challenges outlined above - namely, the lack of standardized processing protocols, the operational complexity of Raman instrumentation in clinical settings, and the variability in data quality - this thesis aims to address these limitations through the design and validation of open-source tools that enhance the reliability and accessibility of Raman spectroscopy. The work is structured around two complementary initiatives : the development of an open-source spectral processing package, and the implementation of the Lumed Raman Acquisition Platform, which integrates custom hardware design with a modular, open-source system control software. The following objectives have been formulated to guide the research and development efforts within these two pillars.

Objective 1 To develop and validate an open-source Raman spectral processing library that delivers improved baseline removal and performance for biological sample analysis, specifically targeting applications in in vivo cancer diagnosis.

- **Sub-Objective 1.1** **Open-Source Raman Processing Library** : To create and disseminate an open-source Python library that provides comprehensive tools for processing Raman spectra obtained from CCD-based spectrometers.
- **Sub-Objective 1.2** **Novel Baseline Removal Algorithm** : To create a novel baseline removal algorithm capable of fitting non-polynomial baselines while ensuring robust performance on biological specimens characterized by low Raman-to-Fluorescence ratios.
- **Sub-Objective 1.3** **Algorithm Performance Evaluation** : To quantitatively evaluate the proposed algorithm's performance against existing methods, de-

monstrating its ability to achieve comparable or superior performance with challenging signals.

Sub-Objective 1.4 **Library Testing and Validation** : To test and validate the processing workflow powered by this library on biological specimen spectral datasets, ensuring robustness, accuracy, and applicability for real-world use cases, particularly in in vivo cancer diagnosis.

Sub-Objective 1.5 **Automatization of processing parameters** : To automatize the optimization of the tuning parameters of the processing algorithms through the use of quantitative metrics, removing remaining human biases.

Objective 2 To develop and validate an integrated hardware-software Raman acquisition platform that ensures rapid setup, consistent measurement quality, and reliable real-time data processing and machine learning-based classification for use in clinical and research environments.

Sub-Objective 2.1 **Sub-Objective 2.1 : Modular Software Architecture** : To design a modular software architecture that enables independent control of key Raman system components (e.g., laser, camera, power meter), facilitating flexible configuration and rapid system deployment.

Sub-Objective 2.2 **Sub-Objective 2.2 : Integration with Open-Source Processing Tools** : To integrate the Open Raman Library into the acquisition software, enabling direct access to data preprocessing, visualization, and spectral quality metrics within the acquisition workflow.

Sub-Objective 2.3 **Sub-Objective 2.3 : Real-Time Processing and Classification** : To implement real-time spectral processing and classification capabilities with minimal latency to support immediate feedback during acquisition.

Sub-Objective 2.4 **Sub-Objective 2.4 : System Validation Using Phantom Models** : To evaluate system performance through controlled measurements on phantom models with known spectral features, ensuring reproducibility and spectral integrity.

Sub-Objective 2.5 **Sub-Objective 2.5 : Evaluation in Simulated Clinical Conditions** : To assess the classification performance and system robustness using biological specimens in simulated clinical settings, with a focus on reproducibility and diagnostic accuracy.

Sub-Objective 2.6 **Sub-Objective 2.6 : Usability and Deployment Efficiency** :

To design a software Graphical User Interface that minimizes setup time and operator input, ensuring usability by clinical personnel with limited technical training and supporting safe, repeatable instrument configurations.

1.3 Thesis Overview

This thesis is organized into five chapters, each addressing a specific aspect of the research undertaken to improve the clinical applicability and methodological standardization of Raman spectroscopy through open-source development. The work is divided into two principal components : the first focuses on the development of a standardized data processing framework, while the second centres on the design and implementation of a fully integrated Raman acquisition platform.

Chapter 2 provides a literature review that reports the scientific foundation of the thesis. It covers the fundamental principles of Raman spectroscopy, the challenges posed by fluorescence interference, and the various computational, instrumental, and sample preparation strategies used throughout the field to address this issue. It also outlines existing applications of Raman spectroscopy in medical diagnostics, with a particular emphasis on cancer detection.

Chapter 3 presents a peer-reviewed publication detailing the development of an open-source Raman spectral processing package. This software suite incorporates all critical pre-processing steps and introduces a novel baseline removal algorithm, called BubbleFill, designed to improve performance in signals characterized by low Raman-to-baseline ratios. The algorithm is validated using synthetic and real-world datasets, demonstrating its effectiveness across a range of sample types and acquisition conditions.

Chapter 4 describes the design and validation of the Lumed Raman Acquisition Platform, which integrates a custom-built hardware system with an open-source control software named the Open Raman Acquisition Software (ORAS). This chapter outlines the technical context motivating the platform's development, details the design of the hardware and the software modules, and presents validation results obtained from phantom and in-human experiments. Emphasis is placed on operational repeatability, real-time feedback, and compatibility with clinical workflows.

Chapter 5 concludes the thesis by summarizing the contributions made to the field, identifying current limitations, and proposing directions for future work. It reflects on the broader implications of open-source approaches in the biomedical optics community and their poten-

tial to facilitate the widespread adoption of Raman technologies in clinical environments.

The thesis also contains three appendices that provide practical documentation to support the deployment of the developed tools. These include a detailed step-by-step assembly guide for the Lumed Raman System, installation instructions for the software platform, and a function reference table for the Arduino-based laser safety control board.

CHAPTER 2 LITERATURE REVIEW

2.1 Intraoperative Tissue Characterization Technologies

2.1.1 Contextualizing Raman Spectroscopy

To fully situate Raman spectroscopy within the current clinical landscape, it is important to examine the principal technologies that have been adopted or proposed for intraoperative tissue characterization. Among these, fluorescence-guided approaches remain the most common, valued for their intuitive visual output and high contrast. The administration of 5-aminolevulinic acid (5-ALA), for example, has become standard practice in high-grade glioma surgery, where it enhances the visualization of malignant tissue and improves resection outcomes [21, 22]. Similarly, indocyanine green (ICG) fluorescence is routinely employed in hepatic, breast, and oral cavity procedures, enabling the intraoperative delineation of tumour margins under near-infrared light [23, 24, 25].

Beyond fluorescence, a range of label-free optical methods provides contrast based on intrinsic tissue properties. Confocal laser endomicroscopy (CLE) and optical coherence tomography (OCT) offer high spatial resolution and are capable of generating histology-like images in real time, making them particularly suitable for soft tissue and lymph node evaluation [26, 27, 28]. Meanwhile, mass spectrometry-based platforms such as the intelligent knife (iKnife) and the MassSpec Pen enable rapid biochemical profiling through lipidomic and metabolomic signatures, and have demonstrated excellent diagnostic performance in a variety of surgical contexts [29, 30].

Anatomical imaging modalities also play an important role. Contrast-enhanced ultrasound (CEUS) contributes vascular and perfusion information in brain and liver tumour resection [31, 32], while intraoperative magnetic resonance imaging (MRI) remains the most comprehensive, though resource-intensive, tool for guiding glioma surgery [33, 34]. Other innovative strategies continue to emerge, including protease-activated fluorescence probes for tumour-specific signal amplification [35], and real-time molecular imaging techniques designed to improve intraoperative detection of residual disease [36]. Reflectance spectroscopy, like Raman, relies on endogenous optical signatures, though with significantly lower biochemical specificity [37].

Taken together, these technologies underscore the breadth of approaches currently employed for intraoperative tissue characterization, each offering distinct advantages in resolution, sensitivity, specificity, or integration into surgical workflows.

2.1.2 Comparative Evaluation of Sensitivity and Specificity

Sensitivity and specificity are key performance metrics for intraoperative tissue characterization technologies, as they directly impact the surgeon's ability to distinguish malignant from healthy tissue and thereby influence margin assessment and re-excision rates. To enhance intraoperative precision and reduce uncertainty at the surgical margin, clinicians now employ a wide gamut of modalities, ranging from label-free spectroscopic tools to contrast-enhanced optical and imaging systems.

Each technique presents unique advantages and trade-offs. For instance, fluorescence imaging modalities benefit from high sensitivity and intuitive visualization, while mass spectrometry excels in biochemical specificity. Technologies like OCT and CLE offer microscopic resolution without the need for staining, and intraoperative MRI provides anatomical context but at a high logistical cost. Devices such as the MarginProbe or protease-activated fluorescent probes aim to streamline decision-making in breast and head and neck surgery, though performance varies considerably.

Table 2.1 summarizes reported sensitivity and specificity values across representative studies, illustrating the variability in diagnostic accuracy among these technologies. It is important to note, however, that direct numerical comparisons between modalities are not always straightforward. Performance metrics such as sensitivity and specificity are often derived using different reference standards, patient populations, tissue types, and definitions of diagnostic success. For example, fluorescence imaging studies may use histopathology of resected tissue as ground truth, whereas spectroscopic or imaging methods might rely on co-localized biopsy validation or intraoperative visual correlation. Furthermore, some studies report per-margin sensitivity, others per-patient, and still others use per-lesion or pixel-based evaluation, each affecting the resulting values. In many cases, device-specific thresholds or algorithmic classifiers introduce additional variability.

TABLE 2.1 Reported sensitivity and specificity for state-of-the-art intraoperative tissue characterization technologies. Values are derived from individual studies, systematic reviews, or meta-analyses in the cited literature ; when precise values were not available, representative ranges were estimated based on contextual descriptions or pooled data across comparable studies.

Modality	**Sensitivity (%)**	**Specificity (%)**	**References**
5-ALA Fluorescence (FGS)	80–100	70–90	[21, 22]
ICG Fluorescence (NIR)	85–95	65–85	[23, 25, 24, 38, 39]
Mass Spectrometry (REIMS, MassSpec Pen)	93–97	93–96	[29, 30]
Optical Coherence Tomography (OCT)	75–90	60–85	[26, 27, 28]
Confocal Laser Endomicroscopy (CLE)	80–95	70–90	[40]
Contrast-Enhanced Ultrasound (CEUS)	70–90	65–80	[32, 31]
Intraoperative MRI	∼95	N/A	[33, 34, 41]
Radiofrequency Spectroscopy (MarginProbe)	70–100	70–87	[42]
Protease-Activated Fluorescence Imaging	85–90	75–85	[35]
Real-Time Molecular Imaging (targeted fluorescence)	86–95	80–89	[36, 43, 44]
Reflectance Spectroscopy	70–90	65–85	[37]

These differences underscore the need for cautious interpretation when comparing diagnostic performance across heterogeneous technologies. While this comparative overview provides a useful benchmark, it remains essential to examine how individual technologies—particularly Raman spectroscopy—perform in actual clinical contexts.

2.2 Clinical Applications of Raman Spectroscopy in Oncology

Raman spectroscopy has demonstrated growing clinical relevance, particularly in oncology, where its label-free biochemical specificity makes it well-suited for intraoperative tissue assessment. Offering non-invasive, high-specificity insights into tissue composition and molecular alterations, it holds strong potential as a diagnostic and surgical guidance tool. Over the

past two decades, numerous studies have validated its utility across a range of cancer types, consistently demonstrating high accuracy in distinguishing malignant from healthy tissue and supporting its integration into clinical workflows.

In breast, linear combination models have been used to fit coefficients of fat and collagen bands as key parameters to attain 94 % sensitivity and 96 % specificity in the classification of cancerous vs normal and benign tissues [45]. In another study, threshold-based classifiers were used to distinguish between benign from malignant pathologies with similar performance [46]. Bands associated with carotenoids, fatty acids, proteins and interfacial water were used to similar ends with 2D Raman microscopy imaging [47]. More recently, a Raman spectroscopy system analogous to the one presented in Chapter 4 was tested with ex-vivo specimen resected during Breast Conserving Surgery (BCS)[48]. The results showed a sensitivity of 93 % and a specificity of 95 % when distinguishing normal breast from ductal or lobular invasive cancer. A detailed analysis of the Raman active bands was provided.

In prostate, multivariate spectral predictive models were used and demonstrated 96 % sensitivity and 91 % specificity for in the identification of malignancy[49]. The same study presented a detailed analysis of Raman peak biomolecular assignment. Deep learning models including Convolutional Neural Network were tested on augmented datasets and reached classification accuracy of (>90 %)[50]. A few years ago, Raman spectroscopy was integrated during brachytherapy procedure [51, 52]. The instrument developed included electro-magnetic trackers for accurate registration in conjunction with the navigation system. This research project illustrated how Raman technology can be integrated with existing diagnostic technologies and provide tangible practical benefits for the patient.

In skin, Raman imaging techniques were used as part of a multimodal diagnostic platform. Ex-vivo tissue samples were fully characterized with microscopy, histopathology, autofluorescence imaging and Raman raster scans. A multivariate model was trained and tested on spectra acquired with a prototype device, showing a sensitivity of 81.8 % and specificity of 92.5 % in the assessment of surgical margins in Mohs micrographic surgery of basal cell carcinoma[53]. In another study, spectra were measured from melanoma, dysplastic nevi, basal cell carcinoma, squamous cell carcinoma and actinic keratosis. Multiple models were developed to distinguish between the different pathology, with varying levels of success [54].

In brain, fibre optics probes have been used intraoperatively to measure Raman signals during surgery. A boosted tree machine learning model was trained on data measured from normal brain, dense cancer and invaded cancer cells with a sensitivity of 93 % and a specificity of 91 %[9]. Another study showed that high wavenumber Raman spectroscopy can detect brain cancer with > 60 % cancer cells in situ [55]. Feature engineering was used to develop a new

representation for Raman data of brain tissue. This adapted analytical method was argued to constitute a crucial step for the broader acceptance of the technology by clinicians. The method was specifically tailored for brain diagnosis and improved interpretability of the Raman signal while retaining model accuracy[19]. A recent multicentre study presented a real-time, in situ brain tumour detection device that could efficiently distinguish between the three most common types of brain tumours. The authors reported diagnostic accuracies of 91 % for glioblastoma, 97 % for brain metastases, and 96 % for meningiomas[56]. This result marked a significant milestone in bringing Raman-based technology a step closer toward the operation room.

2.3 Physical Principles of Raman Spectroscopy

2.3.1 Raman Scattering

The Raman effect is based on the interaction between an incident photon and the molecules of the interrogated sample. When monochromatic light (i.e. from a laser), propagates through a turbid medium, most photons are elastically scattered (Rayleigh scattering)[57]. However, a few photons, typically 10^{-9} to 10^{-6} compared to the intensity of elastic scattering[58], will undergo inelastic scattering which results in a change of energy. This is spontaneous Raman scattering. The change in photon energy corresponds to a vibrational mode from a molecule in the medium. A photon can either lose or gain energy by exciting or exciting a vibration. This is expressed as the presence of either Stokes line or Anti-Stokes line in the measured spectrum. Figure 2.1 presents the jablonski diagrams of the important competing mechanisms that take place with light transport in biological tissue.

2.3.2 Fluorescence

Fluorescence is another light-matter interaction and corresponds to a different inelastic mechanism. With fluorescence, a photon excites a molecule into the vibrational mode of a higher electronic state. After some time, the molecule undergoes non-radiative transitions and relaxes into a lower vibrational level. Eventually, the molecule returns to its original electronic level by emitting a photon of lower energy. Fluorescence imaging is a powerful and widespread tool in medical diagnosis and, as already mentioned, is commonly used for guiding surgery and other procedures[59]. However, the most common techniques rely on fluorophores (fluorecent markers) that are injected intravenously in the patient. The measured signal is that of fluorescent dye such as indocyanine green (ICG), fluorescein isothiocyanate (FITC) or Protoporphyrin IX fluorescence induced from 5-ALA administration instead of being endogenous

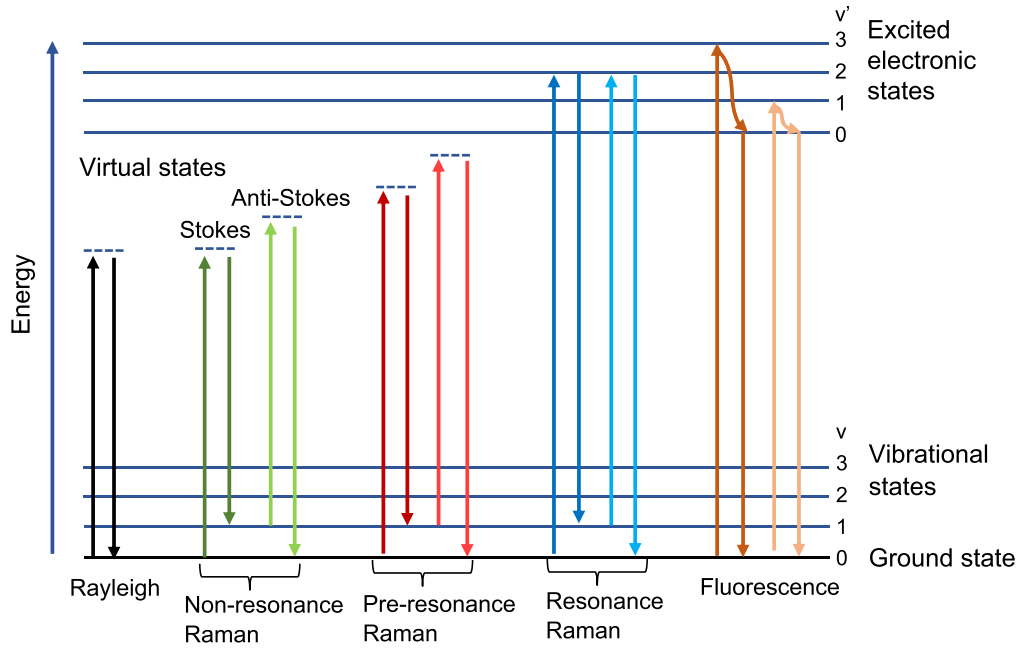


FIGURE 2.1 Jablonski diagram of Rayleigh scattering, Raman scattering and fluorescence. Reproduced from [1]

to the interrogated tissue[60, 61, 62, 63]. Fluorescence Lifetime Imaging (FLIM) and time-resolved fluorescence spectroscopy (TRFS) are more modern variations of the method and have shown promise in measuring signals from fluorophores intrinsic to biological tissue[64]. Yet, the instrumentation necessary is often costly and complex because measurements need to be performed on timescales shorter than the nanosecond. Furthermore, Intrinsic fluorescence endogenous to biological tissues is generally accepted to be less specific compared to vibrational information carried by Stokes and Anti-Stokes transitions while heavily competing with Raman[65, 58, 66, 67].

2.3.3 Measuring Raman Spectra

Spontaneous Raman is the simplest and most common Raman technique for the analysis of biological tissue. Rudimentary clinical systems capable of measuring spontaneous Raman signal safely in hospital environment have been reported since the mid-1990 to early 2000s[2, 68]. Early instrumentation comprised a monochromatic excitation source (i.e. laser), a spectrographic device, a light detector and a fibre optics based instrument (Figure 2.2).

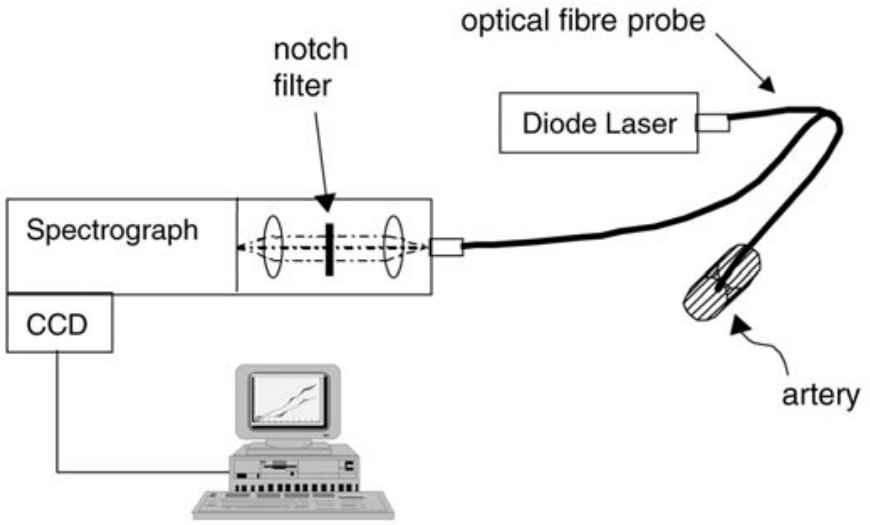


FIGURE 2.2 Raman system from the early 2000s. This system was desigend for rapid data acquisition, portability and safe use in a medical environment. Reproduced from [2]

Over the years, significant advances have been made for each of the components and their integration into this basic acquisition setup (i.e. lasers, spectrometers, cameras and probes). Yet, the general system configuration has seen little change since [9, 69, 70]. Modern Raman acquisition systems still serve a similar function. Light is emitted from an excitation source, guided to the tip of an instrument and illuminates a sample. Backscattered and fluorescence photons are collected by the probe and are guided back to the spectrometer and camera where they are detected. The resulting spectrum is visualized as a histogram plot where the x-axis corresponds to the shift in energy between the incident and detected photons and the y-axis corresponds to light intensity. The energy shift of the x-axis is expressed as wavenumber called the Raman Shift and is calculated as

$$\mathrm{RS}[\mathrm{cm}^{-1}] := 10^7 \left(\frac{1}{\lambda_0[\mathrm{nm}]} - \frac{1}{\lambda[\mathrm{nm}]} \right) \tag{2.1}$$

The literature identifies two distinct spectral regions for the Raman shift : the fingerprint region ($200 - 2000$ cm^{-1}) and the high wavenumber region (> 2000 cm^{-1}). The fingerprint region is rich in biochemical information associated with vibrational bonds of proteins, lipids, nucleic acids and amino acids[2, 71, 10, 72] while the high wavenumber region mostly contains information related to CH-, OH- and NH- stretching modes [73, 74]. Fingerprint is the preferred region when system design limitations force a choice between the two.

The wealth of biomolecular information contained in Raman spectra of biological tissue is

generally too complex and subtle to be efficiently leveraged by humans. Over the past decade, analysis of Raman spectra has instead been relegated to statistical models and Artificial Intelligence (specifically Machine Learning models) for producing biochemical assessments of interrogated samples [75, 76, 77, 13]. These models have been trained to produce fast and accurate medical diagnosis for a wide range of applications detailed in Section 2.2.

2.4 Fluorescence Suppression Strategies in Raman Spectroscopy

The fluorescence scattering cross-section (i.e. quantum efficiency) is significantly higher than that of Raman scattering, leading to substantial interference. This is especially problematic in biological tissue where weak fluorescence can reach quantum yields of $\sim 10^{-4}$ compared to $\sim 10^{-7}$ for Raman [58, 78]. To address this issue, various techniques have been developed to enhance the Raman signal and isolate it from fluorescence interference. These methods can be separated into three categories; computational methods, instrumentation methods and sample preparation methods.

2.4.1 Computational Methods

The research literature reveals several major computational approaches for fluorescence removal in Raman spectroscopy. These methods consist in the use of mathematical (pre)-processing and aim at filtering out the fluorescence contribution from the measured spectra before analysis [79]. One of the earliest methods is to apply a polynomial fit, typically of order 5-6, on the measured data [80]. The polynomial fit will broadly match the slow varying fluorescence contribution while ignoring the sharp and narrow peaks (i.e. Raman contribution) present in the signal. A first improvement on this method was the inclusion of Raman peak rejection from the data fitted with the polynomial curve (i.e. IModPoly) [81]. Traditional polynomial-based methods have evolved to include probability-based adaptive weighting [82] and iterative incorporation of reference spectrum [83]. The use of an algorithm that does not rely on polynomial fit has also become more common. A method that relies on morphological processing was used to remove fluorescence in Raman spectra measured from pigments in works of art [84]. More recently, a new parameter free method that uses Empirical Mode Decomposition was shown to produce comparable results as the IModPoly algorithm without relying on tuning parameters [85].

2.4.2 Instrumentation Methods

Instrumentation methods involve changes to the acquisition system or acquisition workflow and aim at enhancing the strength of the Raman signal or reducing the intensity of fluorescence before or during measurement.

Coherent anti-Stokes Raman scattering (CARS) is one such technique that has shown promise for addressing the fluorescence problem in biological applications [86, 87]. It employs nonlinear optical processes to generate stronger signals than spontaneous Raman scattering. Furthermore, CARS measurements are performed at shorter wavelength than the excitation source which nullifies the presence of fluorescence in the acquisition band, killing two birds with one stone. Unfortunately, CARS imaging remains costly and complex as it requires synchronizing and coupling two high power pulsed lasers. The requirement of a high-energy source can also be a significant constraint limiting medical use in patients.

Fluorescence in a sample can be suppressed by photobleaching before analysis. With this procedure, the sample is illuminated with a continuous exposition to a light source, often of higher energy than the Raman laser. Irradiating the sample in this fashion for an extended period of time leads to the photolytic decomposition of fluorescent molecules [88, 89]. However, it can be slow and cause physical changes, including burning of the samples. It is not suitable for assays involving living organisms or cells due to these drawbacks [90].

Time-gated detection is another method for suppressing fluorescence. It uses ultrafast detectors to discriminate between the instantaneous Raman scattering and the delayed fluorescence emission. This approach relies on short gate widths and high repetition rates to maintain an acceptable detector duty cycle, effectively reducing fluorescence contribution while preserving the integrity of the Raman signal. A study showed a 50- to 80- fold fluorescence background reduction obtained with a fluorescence lifetime of around 4 to 5 ns [91]. However, these benefits come at a high cost. Time-gated systems are complex and require multiple additional optical components including picosecond lasers with high repetition rate, frequency doubling prisms and gateable ICCD systems.

Shifted Excitation Raman Difference Spectroscopy (SERDS) utilizes slight shifts in excitation frequencies to differentiate between Raman and fluorescence [92]. SERDS works by acquiring two spectra in quick succession using two lasers at different excitation frequencies and subtracting the measured results to eliminate fluorescence contributions. This can be performed with multichannel detection to significantly improve signal-to-noise for a given accumulation time[78]. Fluorescence suppression can be further improved by combining the Kerr gating rejection technique with SERDS [93].

In addition to these techniques, researchers have explored various other methods to enhance spectral clarity. For instance, using high numerical aperture lenses in spectroscopic methodologies can suppress detected fluorescence effectively [94]. This approach is applicable across different configurations of Raman spectroscopy, broadening its utility in diverse research scenarios.

2.4.3 Sample Preparation Methods

Sample preparation methods involve interacting with the sample-to-be-measured before the acquisition. Surface-enhanced Raman scattering (SERS) and its derivative are possibly the most important and common methods in this category. With SERS, metallic nanostructures are locally distributed within the sample. The incident laser light excites a plasmon resonance in the metal which can reach an enhancement factor of up to $10^9 - 10^{11}$ of the Raman signal[95, 16, 96, 97]. However, the exact enhancement factor heavily depends on the combination of sample and nanoparticle types.

2.5 The Role of Open-Source in Medical Technologies

The adoption of open-source principles in medical technologies has gained substantial traction in recent years, particularly in the domains of electronic health records (EHR), medical imaging, and clinical decision support. Open-source software (OSS) is defined by its freely available source code, which can be used, modified, and distributed by anyone. In the medical domain, this transparency fosters collaborative development, rapid innovation, and the potential for greater adaptability in diverse clinical settings.

One of the earliest and most notable applications of OSS in healthcare is in electronic medical record systems. Projects such as OpenMRS, GNU Health, and VistA have demonstrated the feasibility and impact of open platforms in low-resource settings and public health systems [98, 99]. These systems are often designed with modular architectures, allowing local customization to meet specific clinical workflows and regional regulatory constraints. Such adaptability has been essential in environments where proprietary systems are cost-prohibitive or poorly suited to local needs.

In medical imaging, open-source tools have been equally transformative. OsiriX, an open source DICOM viewer, stands as a landmark in enabling affordable access to advanced visualization and image processing capabilities [100]. Its success has spurred a broader movement toward community-driven innovation in radiology and diagnostic imaging [101]. These tools allow researchers and clinicians to validate new algorithms, share imaging datasets, and

prototype novel applications—functions that are often restricted by proprietary platforms.

Despite the promise of OSS, its adoption in clinical practice faces several barriers. Concerns about regulatory compliance, data security, user support, and integration with existing hospital IT systems remain prominent [102]. Moreover, sustainability can be a challenge for OSS projects lacking long-term institutional or commercial backing. Nevertheless, initiatives hosted on collaborative platforms such as GitHub have begun to address these issues through better documentation, modularization, and governance models [103].

The use of open-source frameworks is also expanding in clinical laboratory workflows and data management pipelines. Topcu et al. demonstrated how data-centric laboratory systems can be assembled using open-source components, significantly reducing development cost and improving interoperability [104].

Taken together, the open-source paradigm represents a powerful, if still underutilized, driver of innovation in biomedical engineering. Its alignment with principles of reproducibility, transparency, and customization is particularly relevant to the development of next-generation diagnostic platforms—including those based on Raman spectroscopy—where modular design, rapid iteration, and shared validation tools are essential.

CHAPTER 3 ARTICLE 1 : OPEN-SOURCED RAMAN SPECTROSCOPY DATA PROCESSING PACKAGE IMPLEMENTING A BASELINE REMOVAL ALGORITHM VALIDATED FROM MULTIPLE DATASETS ACQUIRED IN HUMAN TISSUE AND BIOFLUIDS

Guillaume Sheehy[a,b], Fabien Picot[a,b], Frédérick Dallaire[a,b], Katherine Ember[a,b], Tien Nguyen[a,b], Kevin Petrecca[c], Dominique Trudel[d,e,f], and Frédéric Leblond[a,b,d,*]

[a]Polytechnique Montréal, Department of Engineering Physics, Montreal, Quebec, Canada
[b]Centre de recherche du Centre hospitalier de l'Université de Montréal, Montreal, Quebec, Canada
[c]McGill University, Montreal Neurological Institute-Hospital, Division of Neuropathology, Department of Pathology, Montreal, Quebec, Canada
[d]Institut du cancer de Montréal, Montreal, Quebec, Canada
[e]Université de Montréal, Department of Pathology and Cellular Biology, Montreal, Quebec, Canada
[f]Centre hospitalier de l'Université de Montréal, Department of Pathology, Montreal, Quebec, Canada

Published in Journal of Biomedical Optics, Feb. 2023

3.1 Abstract

Significance : Standardized data processing approaches are required in the field of bio-Raman spectroscopy to ensure information associated with spectral data acquired by different research groups, and with different systems, can be compared on an equal footing.

Aim : An open-sourced data processing software package was developed, implementing algorithms associated with all steps required to isolate the inelastic scattering component from signals acquired using Raman spectroscopy devices. The package includes a novel morphological baseline removal technique (BubbleFill) that provides increased adaptability to complex baseline shapes when compared to current gold standard techniques. Also incorporated in the package is a versatile tool simulating spectroscopic data with varying levels of Raman signal to background ratios, baselines with different morphologies, and varying levels of stochastic noise.

Results : Application of the BubbleFill technique to simulated data demonstrated superior baseline removal performance when compared to standard algorithms, including iModPoly and MorphBR. The data processing workflow of the open-sourced package was validated in four independent in-human datasets, demonstrating it leads to inter-systems data compatibility.

Conclusions : A new open-sourced spectroscopic data pre-processing package was validated on simulated and real-world in-human data and is now available to researchers and clinicians for the development of new clinical applications using Raman spectroscopy.

3.2 Introduction

Over the last decade, Raman spectroscopy has seen a resurgence in biomedical applications, in good part due to its synergy with emerging advances in data interpretation enabled by recent trends in machine learning and artificial intelligence [105, 11, 10]. The increasing appeal for Raman spectroscopy in medical applications can be traced back to the fact it allows non-destructive (e.g. non-ionizing radiation) interrogation of any biological tissue or fluid, potentially informing on hundreds of biomolecular vibrational bonds within the same measurement. A strength of the technique is that this information can be reinterpreted as a molecular fingerprint lending an interpretation of the material's composition in terms of the relative concentration of proteins and specific amino acids, lipids, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), as well as water and other metabolites [106, 107].

In biomedical sciences, Raman spectroscopy technologies have been deployed at different spatial scales on tissues and on biofluids. Spatial scales consist of confocal microscopy applications for imaging at cellular resolution[108, 109, 14], up to mesoscopic scales for biofluids [7, 17, 110, 16, 3], in situ tissue measurements during surgery [1, 111, 112] and, more recently, at macroscopic scales[113, 114, 115]. For these applications, tissue and fluids were submitted to different pre-processing methodologies. These include in situ in vivo (without sample pre-processing), in situ ex vivo (aqueous solution to maintain tissue viability/integrity), deposition on ex vivo microscope slides for FFPE (formalin fixed, paraffin embedded) tissue, centrifugation for blood and saliva applications. Other important aspects to consider, in all study designs and imaging protocols, relate to the potential impacts associated with measurement temperature[116], heat-induced damages, and signal contamination from preservation chemicals and contaminants.

Raman spectroscopy technology variants include spatial offset Raman spectroscopy (SORS)[117, 118, 119, 120, 121, 122], surface-enhanced Raman spectroscopy (SERS)[123, 95, 96, 124, 125],

and shifted-excitation Raman difference spectroscopy (SERDS)[126, 127, 128]. By modulating the distance between the excitation source and the light re-emission detection locations, SORS allows access to different tissue depths, albeit at the cost of reduced photonic signals, capitalizing on the varying photon sensitivity functions (co-called diffusion "banana shapes") in highly scattering media. SERS uses the plasmonic effect from the coupling of Raman-active molecular bonds with nanostructured metallic surfaces to enhance the inelastic scattering signal of specific molecular bonds, effectively amplifying signals up to several orders of magnitude, although enhancement factors attained in biological material are usually more modest. Finally, SERDS is a method alleviating the need to remove baseline spectral contributions through a direct subtraction of spectra acquired at two closely separated excitation wavelengths. While this technique offers an attractive alternative to the use of background removal algorithms, it does not help resolve the issue of the limited signal-to-background ratio (i.e. inelastic scattering over intrinsic tissue fluorescence) in bio-Raman spectroscopy. Fluorescence from biomolecules usually diminishes monotonically (approximately quadratically) with excitation wavelength (λ), while the inelastic scattering cross-section decreases approximately as λ^{-4}. Excitation wavelengths used in Raman spectroscopy are most often in the red to near-infrared (NIR) range, frequently including 670 nm, 785 nm, 830 nm, and 1064 nm[129]. The latter could, by probing tissue into the short-wavelength infrared (SWIR) region, provide advantages in terms of deeper tissue sensing (reduced elastic scattering and absorption) as well as an improved Raman to fluorescence ratio. However, these advantages need to be balanced against the need to switch sensing technologies, from CCD detectors to InGaAs detection technologies in the SWIR, at the cost of reduced sensor quantum efficiencies and usually less favourable noise characteristics. Time-resolved detection using pulsed lasers and time-gated measurements can also be used to address the signal-to-background problem in biological samples. This is done by limiting light detection to non-resonant interaction phenomena with relaxation time scales occurring on sub-nanosecond timescales, thereby excluding contributions from resonant phenomena such as fluorescence[130].

The common denominator of most of these Raman spectroscopy methods is that they lend spectra composed of up to 1000-2000 intensity bins (e.g. vectors) that require post-acquisition processing data treatment to isolate the inelastic scattering contribution associated with the interrogated material. This is essential because the inelastic scattering (i.e. Raman scattering) contribution is typically orders of magnitude smaller when compared to endogenous fluorescence from tissue biomolecules (e.g. collagen, elastin, NADH and FAD). Moreover, spectra are always distorted by instrument signal contributions (e.g. fluorescence and Raman scattering from optical components), by the spectrally varying response of its optical components, as well as the presence of cosmic rays[131, 132, 133]. Post-acquisition processing

is generally performed as an intermediary step, after spectral acquisition and before (hence pre-processing) visual analysis or machine learning applications (Figure 3.1). In that context, achieving Raman signal pre-processing in a manner that is standardized and that requires minimal expert knowledge is essential for the future development of Raman spectroscopy biomedical applications. This would allow researchers to compare their processed data (i.e. Raman spectra), would enable compatibility and portability of machine learning models developed for similar biomedical applications, and ensure future proofing of currently developed technologies.

This manuscript introduces the Open Raman Processing Library (ORPL), an open-sourced spectral processing python package that includes a new morphological baseline (e.g. fluorescence from bio-molecules) removal algorithm named BubbleFill. This novel algorithm features several advantages over other state-of-the-art methods, such as a reduced reliance on expert knowledge and a decreased risk of under- and over-fitting spectral curves. The library comes with a module allowing to generate benchmark spectra with various Raman, baseline and noise characteristics. The spectra generated from this tool, along with real-world datasets, acquired on different tissue/fluid types using different instruments, are used to demonstrate the use of techniques and the low-level of variability induced by changes in level and shape of the fluorescence baseline and instrument response.

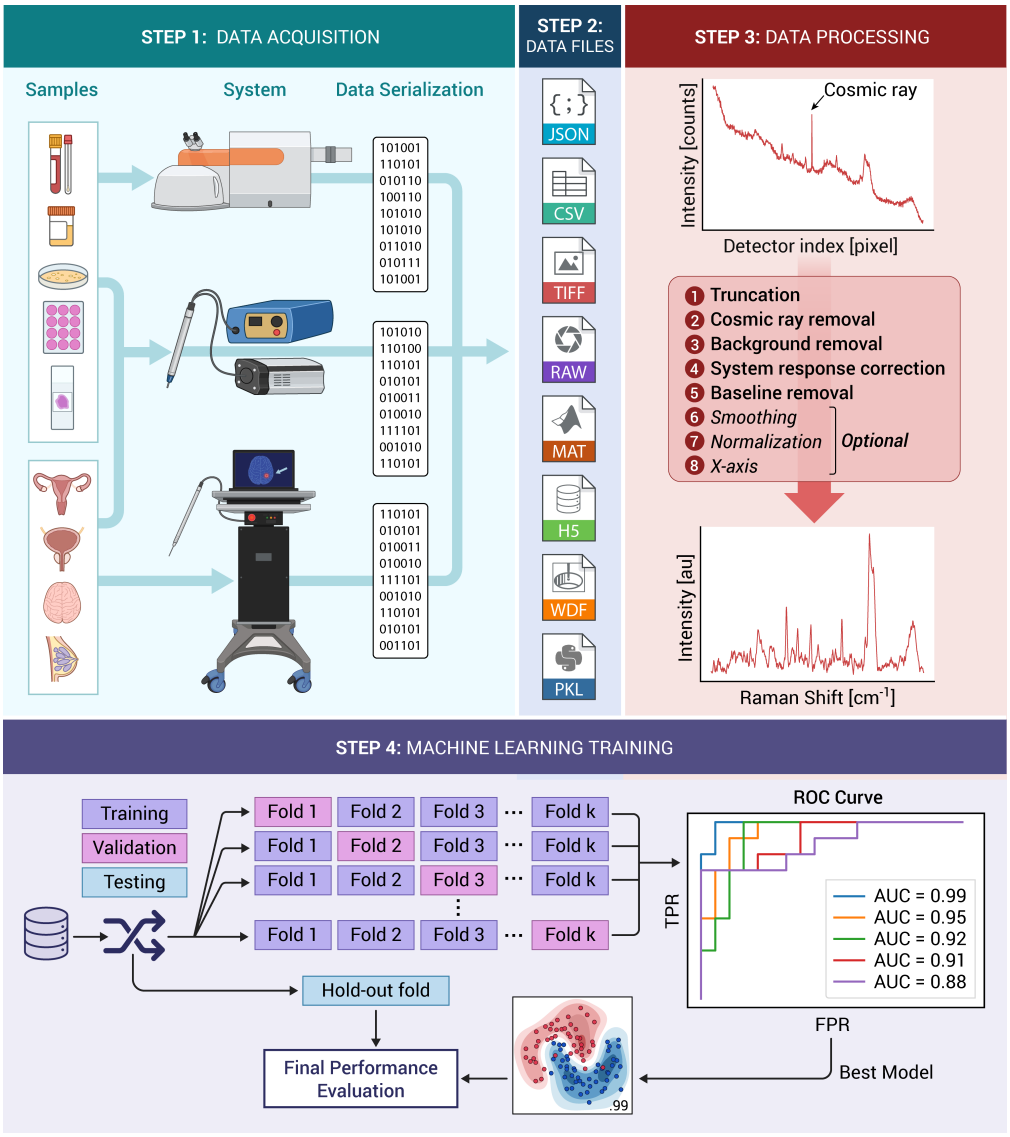


FIGURE 3.1 Depiction of all steps involved in acquiring spectroscopic data using Raman spectroscopy devices and the development of predictive machine learning models. STEP 1 : Spectra are acquired in different contexts (different organs [in situ, ex vivo or fixed] or body fluids) with different instruments, including commercial Raman microscopes, hand-held surgical guidance probes or optical biopsy needles. STEP 2 : Data (Raman spectra and metadata, including patient information) is stored in files with formats that can be processed using readily available tools (e.g. Python, Matlab). STEP 3 illustrates all pre-processing algorithms include in the ORPL open-sourced data processing package. STEP 4 illustrates the process involved in training, validation and testing of machine learning models with performance assessment based on receiver-operating-characteristic (ROC) analysis.

3.3 Methods

The Open Raman Processing Library (ORPL, pronounced "orpel") package offers the necessary tools for processing Raman signals acquired with a variety of different system types, and was optimized to address the specific challenges that arise with biological samples. As there are currently no standards defining which methods should be used for pre-processing steps (both in academia and industry), systems developers tend to include proprietary and obfuscated algorithms and spectral processing techniques within systems controllers and data acquisition software. In some cases, it is not possible nor practical (without extensive software modifications) to export and save a raw spectrum measured by the detector (prior to e.g. cosmic ray removal, averaging, filtering, ins trument response correction). This renders unrealizable the use of a unified processing workflow across Raman datasets measured with different instruments. For this reason, ORPL was designed to be modular, with each module being independent of the others, allowing out of sequence usage.

This section presents the different modules and tools that are available as part of the ORPL package, and the general guidelines for Raman signal pre-processing that our group has developed through multiple studies since 2013 (Figure 3.1.[a]). An example of spectral processing is presented in Figure 3.2, where each intermediary step is illustrated. The acquisition was performed on a sample of Nylon with a spectrometer from the company EMVision (equipped with a Newton CCD camera, Andor) coupled to a 785 nm laser (Innovative Photonics Solutions). Spectra were measured using a handheld probe (EMVision) and the system was operated via a custom in-house software. First, a single spectrum was measured with the excitation source turned off, that is the Background signal (e.g. including ambient light). Then, a series of N spectra was measured with the excitation source turned on, those are the Accumulations or Raw Spectra. Additional measurements were made on an acetaminophen tablet and on a NIST SRM 2241 standard for x-axis and y-axis calibration, respectively. The processing steps are applied in the following order : truncation (section 2.1), Cosmic Ray removal (section 2.2), background removal and combination of the accumulations (section 2.3), y-axis calibration (section 2.4) and baseline removal (section 2.5). Additional steps may include x-axis calibration (section 2.6), smoothing and normalization (section 2.7), but are optional and depend on end-usage of the measured Raman spectrum. Quantitative metrics can also be used to quantify the spectral quality of the spectra in order to allow comparison between spectra acquired with different systems in the scope of different studies (section 2.8).

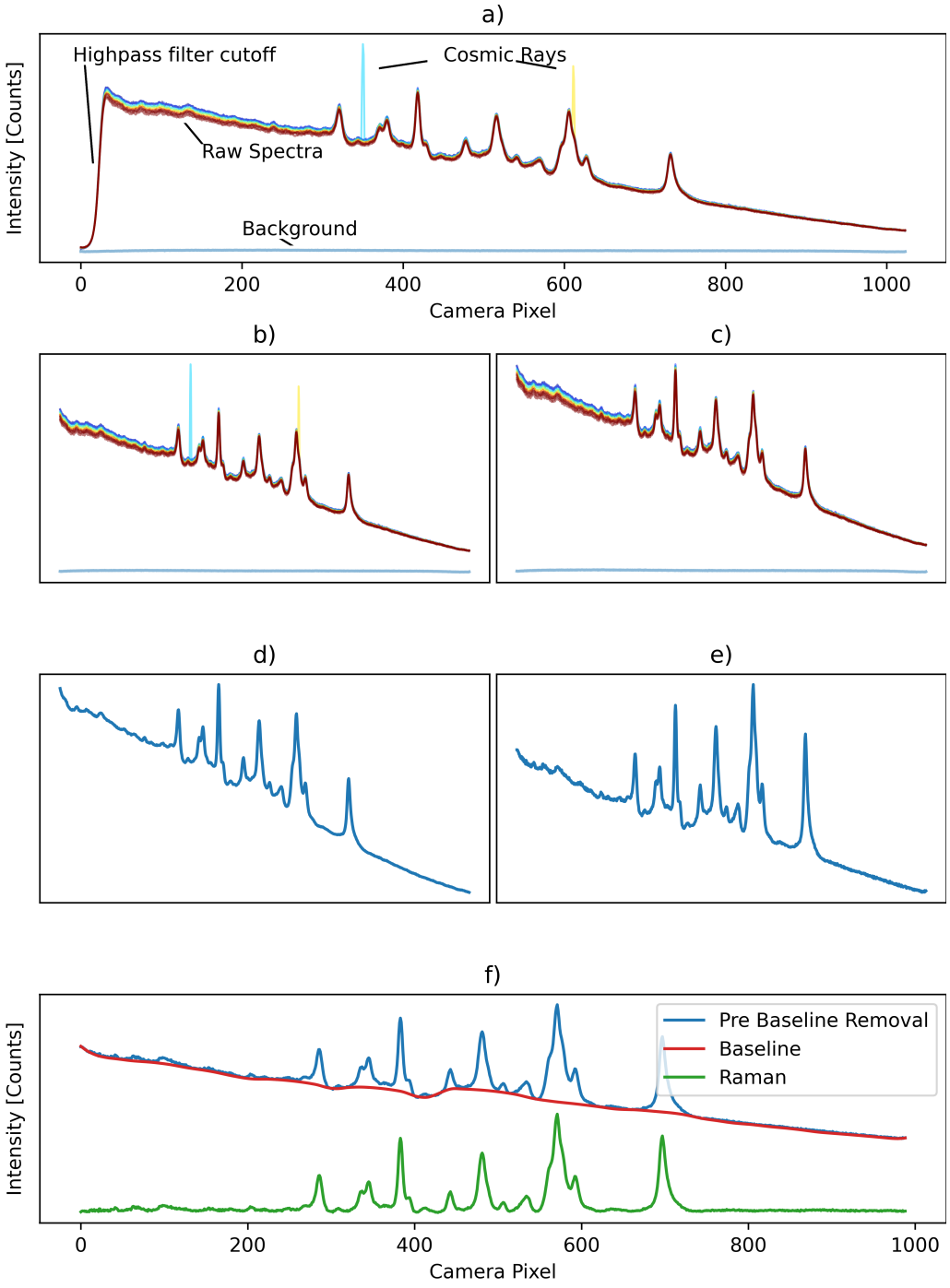


FIGURE 3.2 Overview of processing steps demonstrated on a signal measured from nylon with a point-probe system. a) The raw accumulations and background, b) after truncation, c) after cosmic ray removal, d) after background removal and combining accumulations, e) after y-axis calibration and f) after baseline removal.

3.3.1 Truncation

Raman spectrometers are frequently outfitted with high pass filters (e.g. interference optical filters) to remove Rayleigh-scattered light that is typically orders of magnitude more intense than inelastic scattering. The filter cutoff must be as sharp and as close as possible to the excitation wavelength to limit excitation light bleed-through while maintaining sensitivity to Raman shifts $< 400 - 500$ cm^{-1}. At the same time, the acquisition spectrometer window may extend past the filter cutoff point, toward shorter wavelengths. This results in a spectral shift region at the beginning of every measured signal where the high pass filter transition and the spectrometer window overlap (Figure 3.2.[a]), camera pixels 0-50). This region is truncated and removed from the measured accumulations and background.

3.3.2 Cosmic Ray removal

Cosmic rays randomly hit spectrometer camera pixels during the acquisition of Raman signals, resulting in the appearance of sharp artifacts. Short of reducing exposure time, nothing can be done to mitigate the presence of cosmic rays during acquisitions. Worse, the typically long acquisition times required for biological samples (seconds, up to minutes in some cases) makes the presence of cosmic ray artifacts likely. Fortunately, they are relatively easy to remove using one of the following methods.

The first approach, implemented in ORPL's `crfilter_single()` function, relies on the localized nature of the cosmic ray artifact. Since their spectral span is usually limited to 3-5 camera pixels, it is possible to 1) use the numerical derivative of the spectrum, 2) identify cosmic ray artifacts using an adaptive threshold and 3) remove the artifact from the original spectrum using interpolation. This method should be limited to cases where it is impossible to acquire several spectra during an acquisition (i.e. a single accumulation, N = 1), or for removal of cosmic rays from a background signal. This is because it can be difficult to tune the algorithm parameters to remove every cosmic rays while keeping the signal of interest intact.

The second approach, implemented in ORPL's `crfilter_multi()` function, relies on the random nature of cosmic rays. Because it is exceedingly unlikely that two spectra or accumulations exhibit a cosmic ray artifact at the same wavelength (same camera pixel), it is possible to 1) compare intensity across spectra (from different accumulations) wavelength per wavelength to compute the disparity, 2) identify outliers as artifacts, and 3) replace the region neighbouring the artifact using interpolation. Related spectra can be the different accumulations from a single site using a point-probe system, or spectra acquired over a small sample

region using an imaging system such as a Raman microscope.

In general, the second method should be prioritized whenever possible, as it tends to identify cosmic rays more reliably while minimally affecting the filtered signals and is easier to tune. However, it can only be used effectively on multiple signals or accumulations, which are not always available.

3.3.3 Background removal and combining accumulations

Backgrounds are measured before an acquisition to account for contributions to the signal that are not related to the sample, such as ambient light. They should be measured with the same experimental parameters (e.g. exposure time and instrument position whenever possible) as the main acquisitions, but with the excitation source turned off. It is also important to make sure no cosmic ray artifacts are present on a background spectrum to avoid introducing a downward spike in the signal. This can be done by using the `SSCR_filter()` on the background before removal when a single background signal is measured, or with `MSCR_filter()` otherwise. After this, the background is directly subtracted from the spectrum, after normalizing for exposure time. When multiple accumulations are measured for the same acquisition site, they are combined into a unique spectrum by computing their arithmetic mean, leading to improved signal-to-noise ratio (SNR). If that is the case, the background should be subtracted after the accumulation average to avoid unnecessary computational steps.

3.3.4 y-axis calibration

Instrument response correction is a necessary step for any spectroscopic measurement. Typical methods involve measuring using a spectrum from a reference calibration light source of known emission profile to recover the instrument response function (IRF). However, this approach can be challenging to integrate in most Raman spectroscopy acquisition workflows due to the requirement of additional equipment and the difficulty of positioning the calibration lamp[20]. Instead, an alternate method for calibrating Raman instruments (e.g. correct for filter, detector etaloning and quantum efficiency effects[134]) is based on measurement on a standard reference material (SRM) calibrated and manufactured by NIST (SRM-2241 for 785 nm excitation)[129]. Because this method does not involve the use of additional instruments (irradiance source), it is better suited for use in clinical environments where time and space are limited. Additionally, reference materials can be encased in a custom 3D-printed enclosure to facilitate positioning of probes or instruments and improve systems calibration repeatability. With this method, the IRF is computed from a Raman spectrum measured

on the SRM using experimental parameters (laser power, exposure time) that maximize the sensor's dynamic range. Then, the measured spectrum is processed using the steps previously described : truncation, cosmic ray removal, background removal. After those steps, the instrument's IRF is computed from the known theoretical fluorescence response ($SRM_{theoretical}$), given as polynomial coefficients by the manufacturer and the measured signal ($SRM_{measured}$) as

$$IRF = \frac{SRM_{measured}}{SRM_{theoretical}} \tag{3.1}$$

The instrument response is then corrected for by dividing the Raman spectrum acquired by the IRF vector.

3.3.5 Baseline removal

Standard algorithms

Baseline removal is the most difficult yet most critical step in the processing of a Raman spectrum. Because the probability of Raman scattering is orders of magnitude smaller compared to elastic scattering and fluorescence (in biological materials), even small artifacts introduced during this step can have a disastrous effect and completely overshadow the Raman signal. A common baseline removal method consists in using a polynomial fit (typically of order 5 or 6) coupled with a peak rejection rule (e.g. iModPoly algorithm[81]). The peak rejection rule is used to exclude regions of the signal that feature Raman peaks. Ideally, the entire Raman signal is excluded, leaving only a smooth baseline to be fitted with the polynomial function. This method is simple and can be effective in cases where Raman peaks are easily detected with simple peak finding algorithms and if the baseline can be well modelled by a polynomial function. However, it is not often the case for biological samples. Although intrinsic fluorescence can be adequately modelled as a polynomial function, the same cannot be said for absorption and scattering within the visible-IR range. Furthermore, biological Raman peaks are weak and complex, which makes them far harder to automatically detect compared to pure chemical compounds or inorganic materials. The result is that polynomial methods for baseline removal are difficult to tune and often introduce important artifacts when used for spectra measured on biological samples (Figure 3.3.[a]). A method developed by Perez-Pueyo based on morphological processing (referred to as MorphBR moving forward) has shown to be better suited than polynomial fitting for biological applications[84]. However, baselines removed using this algorithm feature a "jagged staircase effect" and a "hill", that is as wide as the filtering window, near the beginning of the spectrum (Figure 3.3.[b]).

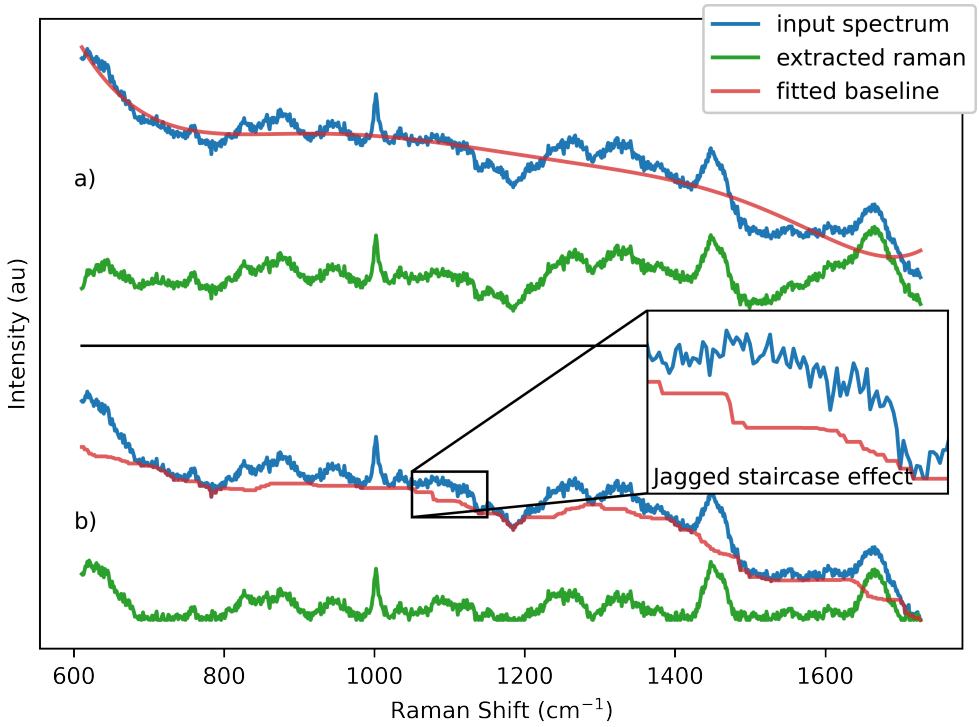


FIGURE 3.3 Demonstration of problems with common baseline removal algorithms. The spectrum used in this example was measured with a Renishaw Raman microscope on saliva samples[3]. a) iModPoly (polynomial fit), b) MorphBR (morphological baseline removal).

BubbleFill algorithm

Inspired by the approach of Perez-Pueyo (MorphBR), we developed a new algorithm based on morphological processing that results in a smoother baseline fit, introduces fewer artifacts and for which fitting parameters can be tuned to different levels across the x-axis. We named this new algorithm BubbleFill because the fitting process uses circular bubbles to fill the region underneath the spectrum. A detailed flowchart of the algorithm is presented in Figure 3.4 and an illustration of the bubble growth loop is presented in Figure 3.5. First, the overall slope of the spectrum is removed using a linear fit and the result is scaled for the x-axis and y-axis to span the same range (square aspect ratio). The baseline estimate is initialized at 0 over the entire x-axis. Then, circular bubbles are iteratively grown under the scaled spectrum, starting with a bubble of diameter equal to the spectrum's length aligned on the center of the x-axis. The bubble "pops" when it reaches the spectrum, and two new bubbles start to grow on each side of the contacting point. A bubble's diameter and alignment depend on the x-axis region where it is grown. After every bubble has popped, the new baseline estimate is updated as the maximum value between the bubble and the current baseline. This process repeats until

every new bubble has reached the critical minimum diameter specified as a tuning parameter. Finally, the baseline estimate is smoothed using a Savitzky-Golay filter and scaled back to the original spectrum y-axis range. The final Raman signal is then obtained by the subtraction of the baseline estimate to the original spectrum. The only tuning parameter corresponds to the smallest allowed bubble diameter. The smaller they are allowed to grow, the more aggressive is the baseline fitting process and vice versa. Additionally, it is possible to specify different bubble diameters across the x-axis, effectively achieving a different degree of sensitivity over different regions of the input signal.

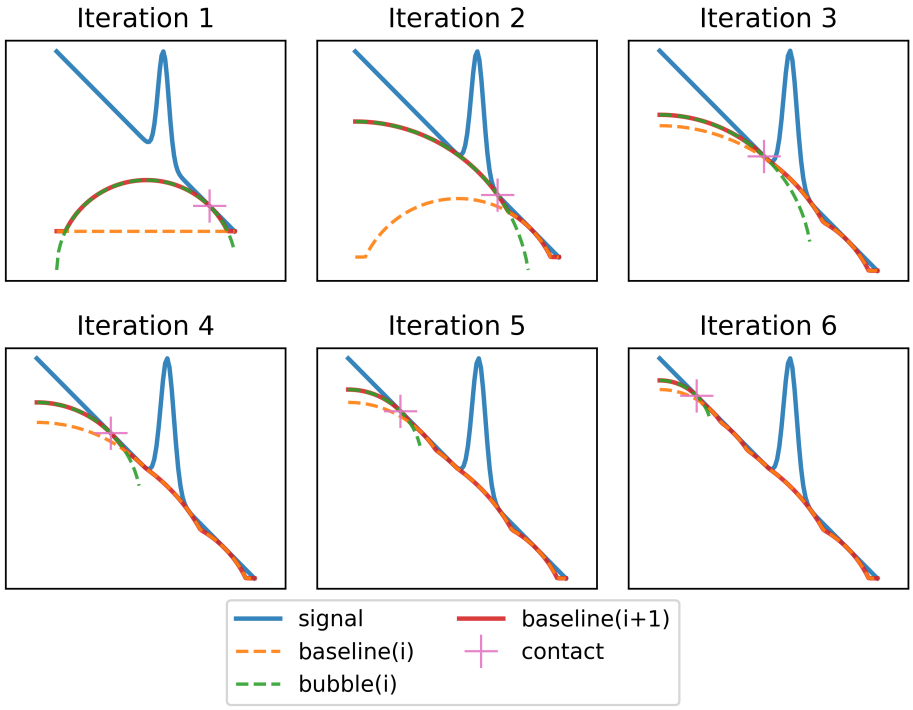


FIGURE 3.5 First 6 iterations of the bubble growth and baseline update iterative process that is at the core of the BubbleFill algorithm. Only the first 6 iterations are displayed, resulting in a small, obtuse angle toward the right side of the peak. Additional iteration reveals that a small bubble is eventually grown in that region and significantly improve the baseline fit. The Savitzky-Golay filter applied to the baseline fit after the bubble growth process further smooths possible remaining sharp angles.

x-axis calibration

For the same reason that the y-axis calibration of a Raman instrument should ideally not be made using a calibration light source, the x-axis Raman shift needs to be computed from a reference sample's spectrum. Reference samples used for this purpose should have a Raman

to Baseline Ratios (RBR), that is the ratio between the intensity of the tallest Raman peak to the maximum of the baseline signal, of at least 0.2. Additionally, it is preferable to use samples featuring narrow Raman peaks that can be easily identified with an automatic peak finding tool, such as SciPy's `find_peaks()` function. Acetaminophen and Nylon are two reference samples well suited for this purpose. They both feature narrow peaks that are uniformly spread over the $0 - 2000$ cm^{-1} region, have a Raman to baseline ratio above 0.5, are shelf stable while not requiring tedious maintenance and can easily be brought into sterile environments.

To calibrate a Raman spectrometer's x-axis, a spectrum is measured from a chosen reference sample, then the spectrum is processed using the steps described in Section 2.1 to 2.5. The camera pixel location index of the most prominent Raman peaks are identified using a peak finding tool (e.g. SciPy's `find_peaks()` function). Finally, a polynomial fit of order 2 or 3 is used to create a map conversion between camera pixel index and Raman shifts. An important note, when calibrating the x-axis for machine learning applications, is that it is critical to interpolate every spectrum over a common x-axis. Otherwise, there is no guarantee that all spectra are expressed in the same vector space or, in other words, that the intensity vectors across spectra correspond to the same wavelength/energy.

Smoothing and normalization

Smoothing and normalization are highly dependent on the end use of a Raman spectrum. When spectra are intended for visual analysis, they are smoothed using an average moving filter and normalized so that the resulting spectrum's minimum is 0 and maximum, or intensity at a chosen Raman shift, is 1. This type of normalization helps with visual comparison and assessment of spectra acquired on samples that cover a wide range of absorption and scattering. When spectra are used in machine learning applications, however, they are not smoothed and are normalized using the Standard Normal Variate (SNV) method. The SNV normalized ($s^\star$) transformation of a signal (s) is given as

$$s^\star_{i\in[0,N]} := \frac{s_i - \bar{s}}{\frac{1}{N}\sum_i (s_i - \bar{s})^2} \tag{3.2}$$

where the normalized signal has a mean of 0 and a standard deviation of 1.

3.3.6 Spectrum quality factor

The definition of a general spectra quality factor is necessary for comparison and ranking of signals from different datasets. However, common metrics such as the Signal-to-noise ratio (SNR) or the Raman-to-Baseline ratio (RBR) tend to be highly dependent on the instrument (spectrometer slit and resolution, detector efficiency) and acquisition parameters (excitation power, exposure time). Because of this, Raman signals measured from different sample types or using different systems tend to have vastly different SNR and RBR. Often, a top-quality spectrum from one dataset would have an SNR and RBR smaller than every spectrum from another dataset. It is therefore impracticable to use either metric for general spectrum quality comparison across multiple datasets or instruments. Instead, the quality of the Raman spectra presented in this work was measured using the Average Signed Squared Intensity (ASSI) as defined by

$$ASSI := \frac{1}{N} \sum_{i=1}^{N} \operatorname{sgn}(r_i^\star) \cdot r_i^{\star 2} \tag{3.3}$$

where $r^\star$ is an SNV normalized Raman spectrum and $\operatorname{sgn}(x)$ is the sign function of x, that is -1 or 1 whether x is negative or positive. Given this definition, the ASSI of an arbitrary Raman signal is bound between -1 and 1. Because an SNV transformed signal's average is 0, squaring the Raman intensity is necessary for the ASSI computation sum to return a non-zero value. Additionally, this non-linear scaling and the use of the sign function favours large intensity peaks and penalizes signals that have intensity drops below the signal average. In summary, if a signal contains few large and narrow peaks, its ASSI will be large, whereas if a signal contains many small and broad peaks, its ASSI will be small. Finally, a signal that contains only stochastic noise will have an ASSI of 0.

3.4 Results and discussion

Two different complementary approaches were used for the validation of ORPL. First, a module for the numerical generation of synthetic spectra was implemented as part of the library. This module enables the creation of synthetic benchmark spectra that can be used for quantitative testing of the novel BubbleFill baseline removal algorithm. Then, experimental datasets from previous studies have been compiled and uniformized in a single dataset. The uniformization consisted in the conversion of all data files into the `.json` open standard file format and the bundling of acquisition metadata as object properties for each spectrum. The compound dataset was used for validation of the capabilities of ORPL for the processing

of spectra from different biological sample types (including different tissue types or similar tissues with different sample preparation) across different systems. Here, validation served to confirm that the BubbleFill algorithm addresses current limitations of other baseline removal algorithms and that the processing workflow implemented in ORPL enables the recovery of Raman spectra featuring vibrational bands commonly expected in biological samples.

3.4.1 Comparison of baseline removal algorithms : synthetic tool

Synthetic spectra (S_i) were modelled as a combination of three signal components : Raman (R_i), baseline (B_i) and noise (N_i). Generation of benchmark signals in this manner allows a fine control on the desired signal-to-background ratio (SBR) and SNR. Additionally, outputs given by a baseline removal algorithm can be compared to the original signal components with which the input spectrum was generated to compute overall fitting error. However, the Raman and baseline components used for the generation of synthetic spectrum need to be independent of the baseline removal algorithm to be tested to limit possible biases. The formula used to generate synthetic spectra is :

$$S_i = \frac{sbr \cdot R_i + B_i}{\max(sbr \cdot R_i + B_i)} + N_i \tag{3.4}$$

where the index i runs from 1 to N, that later being the total number of bins composing each spectrum.

The Raman component used for generation of benchmark spectra were experimentally measured on acetaminophen, Nylon and PDMS samples using a point probe system[135] and were processed using the aforementioned workflow. The experimental Raman spectra were hand-fitted as a superposition of Gaussian curves (Figure 3.6.[a]) to ensure only clean Raman peaks remained, and to limit eventual biases introduced by the baseline removal algorithm. For the baseline component, spectra were measured on aluminium and nigrosin. As neither of the signals featured noticeable Raman peaks, it was concluded that their spectral responses consisted essentially of pure fluorescence (baseline) signals. These experimental baselines were smoothed using a 50-pixel wide average moving filter to further remove residual traces of noise resulting in the Baseline components (B_i) used for the generation of benchmark spectra. Together, the two baseline components and three Raman components (Figure 3.2.[b]) enable the creation of a wide variety of possible synthetic spectrum suited for use as quantitative benchmark tests for baseline removal algorithm validation and optimization (Figure 3.6.[c]). Finally, the noise component (N_i) is generated following a normal distribution with an average of 0 and a specified standard deviation.

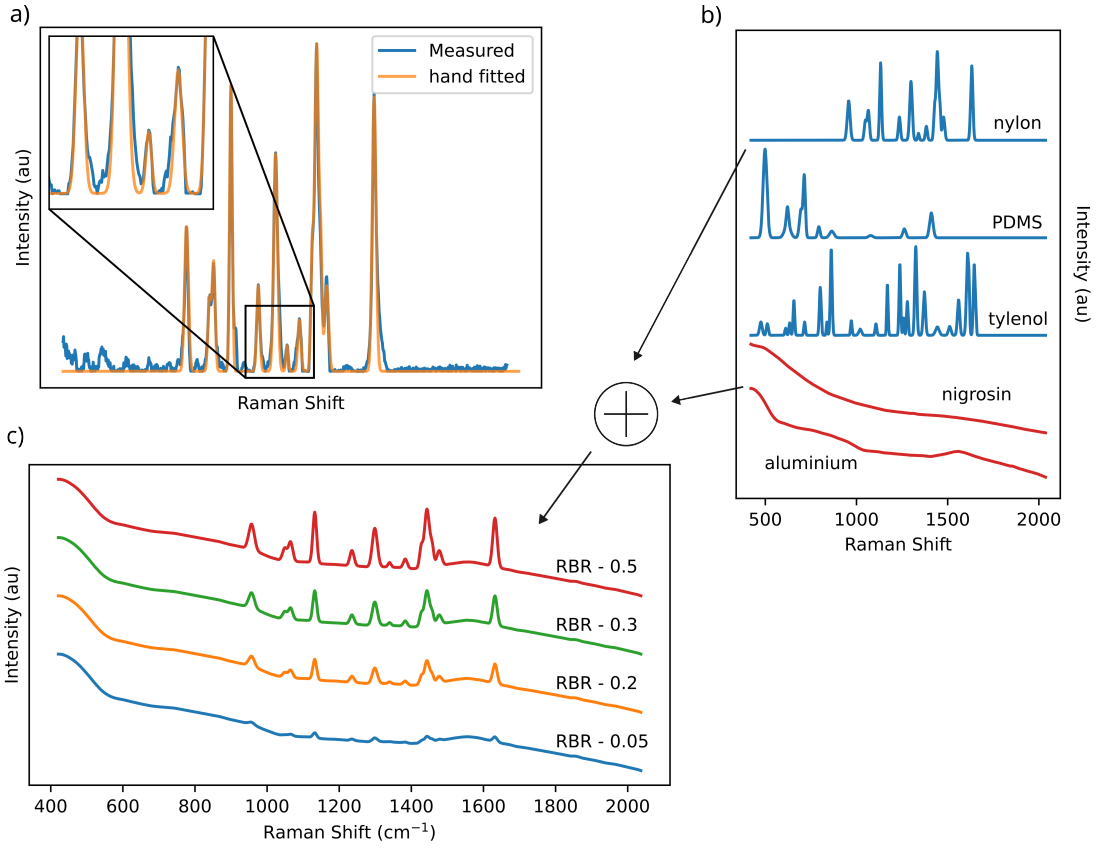


FIGURE 3.6 Baseline algorithm benchmark with synthetic spectrum generated from the aluminium baseline and acetaminophen Raman with a Raman to baseline ratio of 0.05.

The metric chosen to evaluate the performance of the different baseline removal algorithms implemented in ORPL is the normalized Mean Squared Error ($nMSE$). The $nMSE$ between a Raman signal computed from a baseline removal algorithm (R_i^c) and the Raman target (R_i^t) used in the generation of a benchmark signal is defined as

$$n\text{MSE} := \frac{\sum_i (R_i^t - R_i^c)^2}{(\sum_i R_i^t)^2} \tag{3.5}$$

3.4.2 Baseline removal benchmark without noise

A benchmark spectrum is generated using the acetaminophen Raman and aluminium baseline components with an SBR of 0.05. The spectrum's baseline is removed with the BubbleFill, MorphBR and iModPoly algorithms and their respective Raman outputs are compared to the target. The results presented are the optimal fit of each algorithm (smallest $nMSE$) obtained by sweeping the possible range of value for their respective tuning parameter. This method

is used to compare baseline removal algorithms in a 'best-case scenario' levelled playing field.

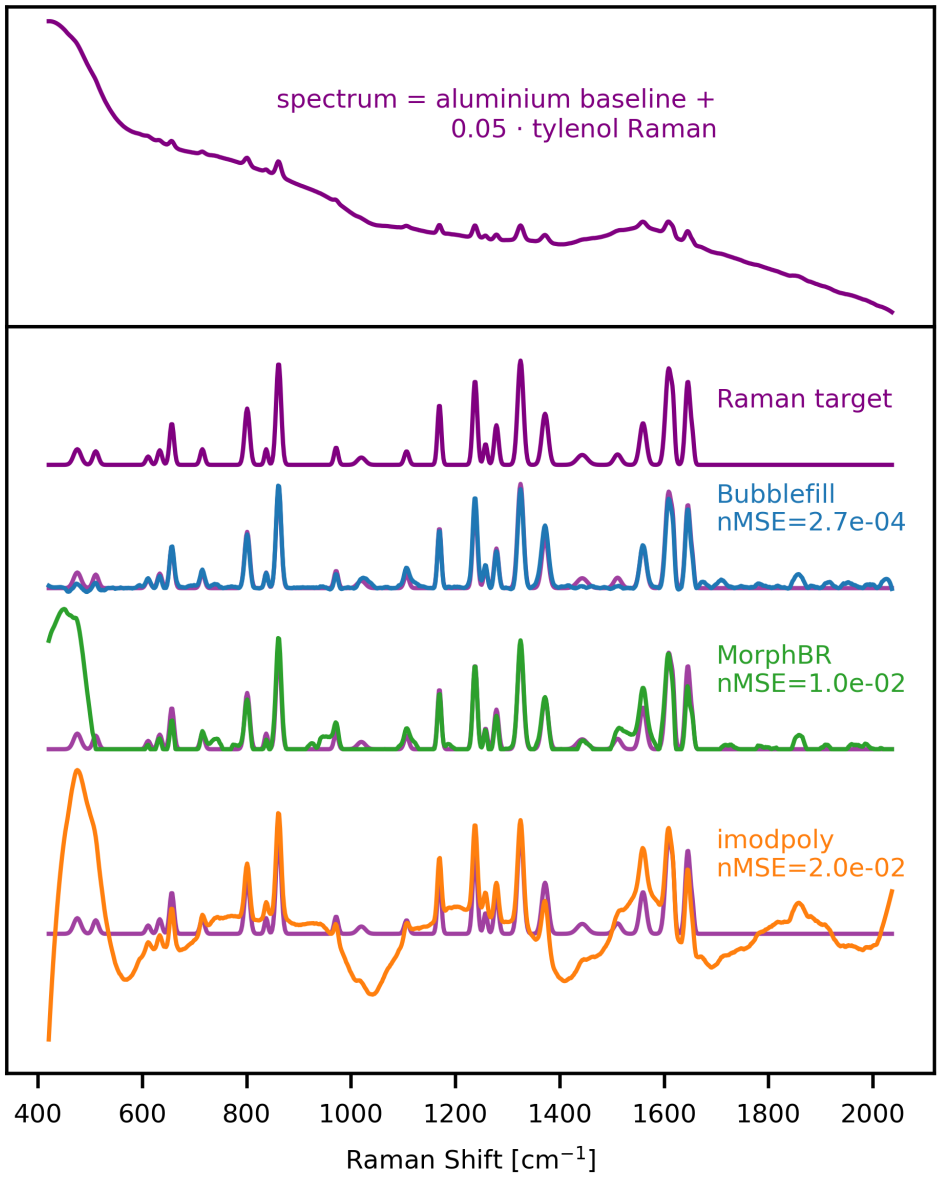


FIGURE 3.7 Baseline algorithm benchmark with synthetic spectrum generated from the aluminium baseline and acetaminophen Raman. Signal-to-noise : 0.05. Noise standard deviation : 0.

This first example highlights the shortcomings of polynomial fitting methods when removing baselines that feature "localized bumps", such as the ones seen in many biological samples (Figure 3.7). Even though not all biological samples feature a baseline as difficult to remove as aluminium, localized bumps can still be introduced in the measured signal when

acquisitions are made over an aluminium substrate. This is relevant and problematic because aluminium slides are becoming an increasingly popular and cheaper alternative to calcium fluoride (CaF2) slides for Raman microscopy [110, 134, 136]. The Raman spectrum given by the BubbleFill algorithm closely matches the target except for the two smaller peaks near $1500\,\mathrm{cm}^{-1}$. In most instances, using a different tuning (e.g. larger bubbles) would result in fewer small peaks being removed at the cost of a worse baseline fit. This is the general trade-off when tuning any baseline removal algorithm. However, with BubbleFill, it is possible to use a different tuning for different regions of the spectrum. For instance, the bubbles grown in the 1400 to 1600 cm^{-1} region could be much larger than for the rest of the spectrum. This would result in a similar baseline fit, but the small peaks near 1500 cm^{-1} would be preserved. This feature is, to our knowledge, unique to BubbleFill and could lead to significant improvements to baseline removal in some applications. Furthermore, additional testing with different combinations of baseline and Raman components and for different signal-to-background ratios indicate that BubbleFill outperforms other tested algorithms in numerous instances (Figure 3.8). The largest difference in performance was observed for spectra generated with the aluminium baseline and acetaminophen or nylon Raman components. In these cases, BubbleFill has an nMSE nearly two orders of magnitude smaller than the second best performing algorithm. It fell behind only for spectra generated with nigrosin baseline and nylon or PDMS Raman components, where it still remained the best option when the Raman to baseline approached 0.

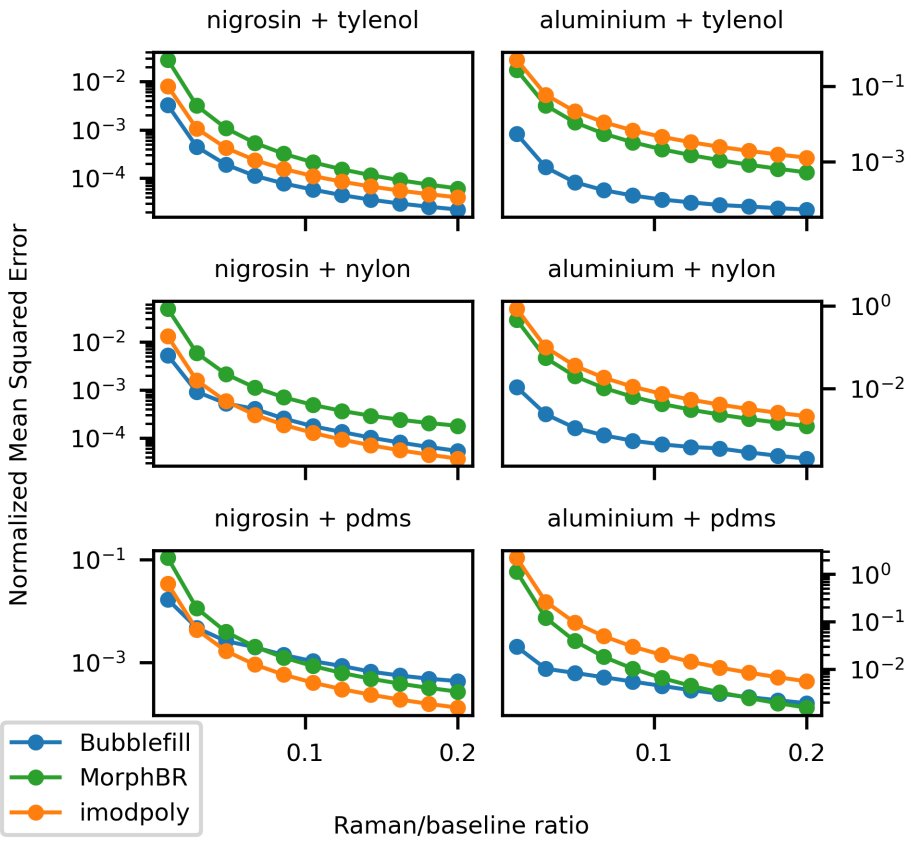


FIGURE 3.8 Normalized Mean Squared Error as a function of the Raman to baseline ratio for the three algorithms tested. Each graph is for spectra generated with a different combination of baseline and Raman signal components.

The different algorithms tested in this work along with the results presented in this section are summarized in Table 3.1.

TABLE 3.1 Baseline algorithm comparison summary. Execution time was measured on a benchmark spectrum generated with the Nylon Raman, Aluminium baseline, Raman/baseline ratio of 0.5 and noise standard deviation of 0.1.

Algorithm tested	General Performance	Tuning parameters	Intuitiveness of tuning parameters	Noticeable shortcomings	Execution time on a signal of size 1000
iModPoly	–	Order of polynomial fit, precision target	–	Fails to fit non-polynomial baselines.	2.6 ms ± 7.65 µs
MorphBR	+	Morphological filter window size	–	"Jagged staircase" effect in the fitted baseline. Introduces a "hill" in the output Raman near the origin of the x-axis.	727 µs ± 1.71 µs
BubbleFill	++	Minimal bubble width	+		743 µs ± 2.44 µs

3.4.3 Baseline removal benchmark with noise

Baseline removal algorithms have been tested on spectra with noise to confirm that the results shown thus far can translate to real-world applications. However, the method of comparing algorithms' performance via the mean squared error becomes impractical when noise is added to the input signal. The problem is that the MSE between an algorithm's computed Raman and the target is overwhelmingly correlated with the added noise itself. This makes the computed MSE seemingly identical between the three algorithms in every scenario tested, even when differences remained visually noticeable in the respective outputs. Instead of comparing nMSE across the different algorithms on noisy spectra, two examples have been chosen to illustrate tendencies also observed in the experimental datasets presented in the following section. In both examples, the input signal was generated using baseline and Raman components as described before, but with the addition of noise. The noise signal added follows a normal distribution with average 0 and standard deviation of 0.01 - that is 1 % of the baseline's maximum.

Figure 3.9 shows the baseline removal benchmark on a spectrum generated from nigrosin and nylon with Raman-to-noise ratio of 0.15. This example was chosen because it is a best-case scenario for the iModPoly algorithm - the nigrosin baseline is smooth, and the nylon peaks are easily identifiable and located toward the centre of the spectrum. Yet, both BubbleFill and MorphBR performed similarly to iModPoly.

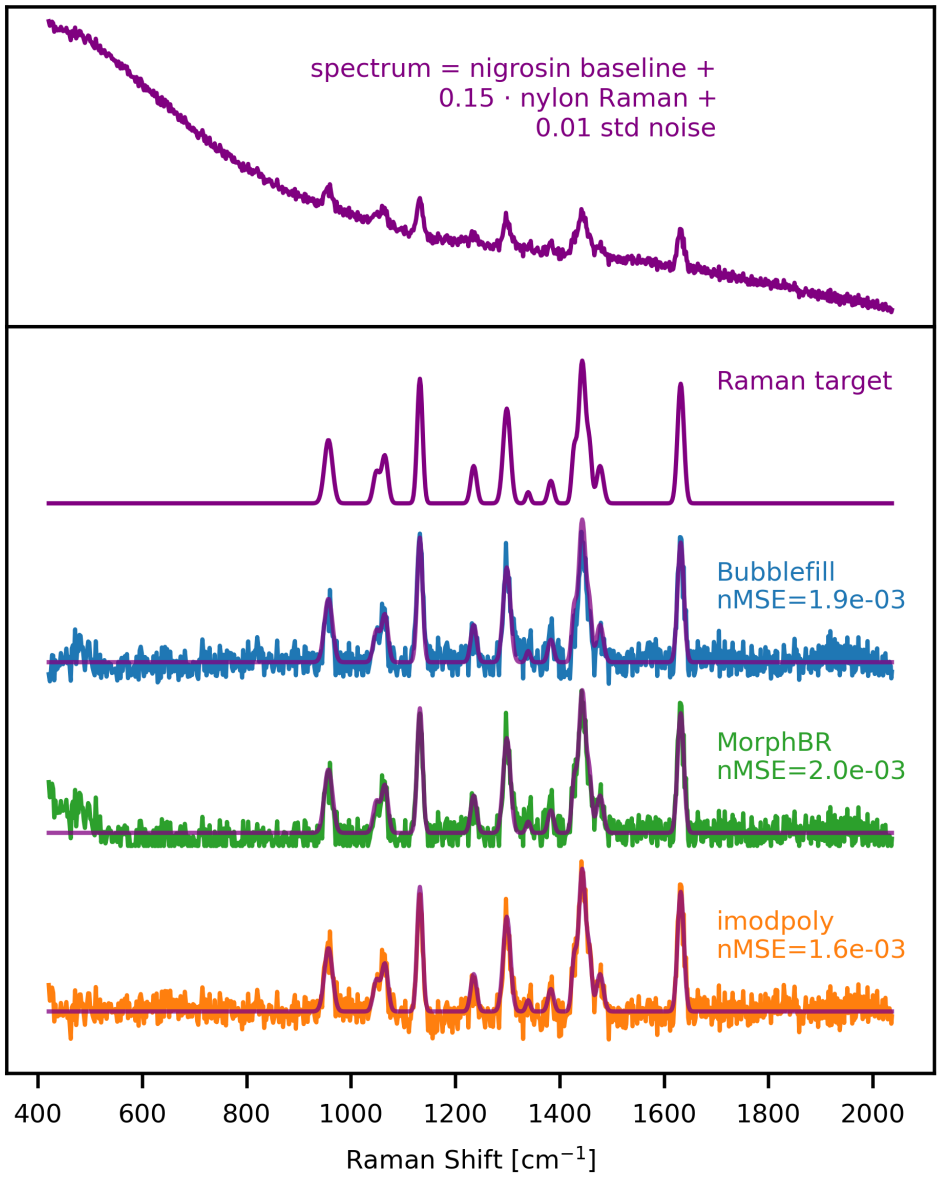


FIGURE 3.9 Baseline algorithm benchmark with synthetic spectrum generated from the nigrosin baseline and nylon Raman. Raman to baseline ratio : 0.15. Noise standard deviation : 0.01.

Figure 3.10 shows the baseline removal benchmark on a spectrum generated from aluminium and PDMS. Similarly to the tests performed on signals without noise, iModPoly failed to correctly remove the aluminium baseline while BubbleFill and MorpthBR managed a near perfect recovery of the Raman target except for a small discrepancy (shared by every algorithm) for the first peak near 500 cm^{-1}. In general, none of the algorithm's performance were significantly affected by the addition of noise to the input spectrum.

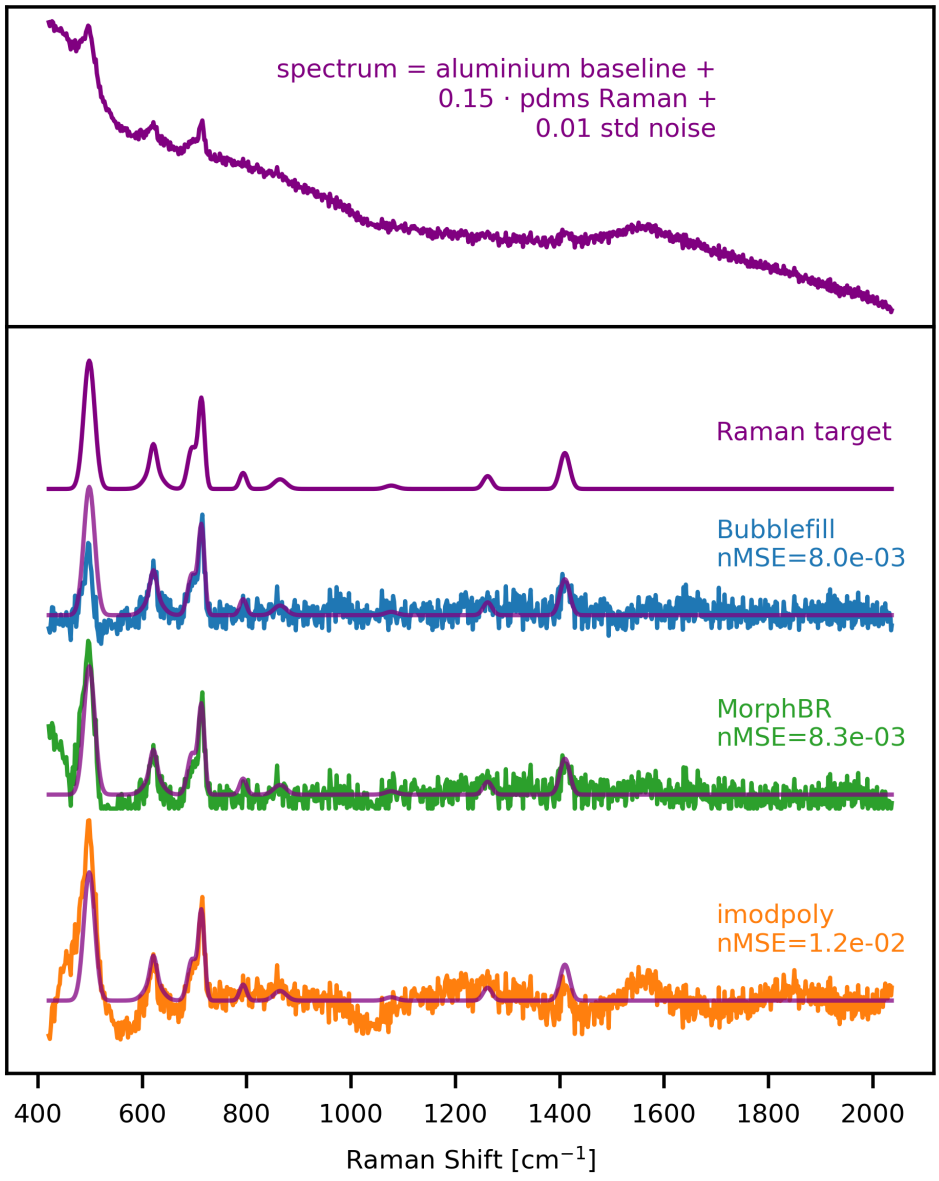


FIGURE 3.10 Baseline algorithm benchmark with synthetic spectrum generated from the aluminum baseline and PDMS Raman. Raman to baseline ratio : 0.15. Noise standard deviation : 0.01.

3.4.4 Testing ORPL with different experimental datasets

We tested the capability of ORPL on four datasets acquired with three different Raman platforms to cover a wide range of sample types, instruments and sample preparation methods. The first dataset (1719 spectra) consists of in vivo brain tissue spectra measured with the hand-held probe from Reveal Surgical and provides a good reference for spectra acquired in

a surgical workflow. The second dataset (524 spectra) consists of in vivo and ex vivo prostate tissue spectra measured with a custom lab-built system using a commercial EmVision LLC handheld probe. The third (8670 spectra) and forth datasets (7774 spectra) consist of paraffin fixed prostate tissue slices and dried saliva samples respectively and were measured with a commercial Renishaw Raman microscope as 3D Raman maps. Combined, the four datasets amount to a total of 18,687 individual acquisitions (not counting repeated accumulations) that cover a wide range of signal-to-background and signal-to-noise ratios as confirmed by the spectral quality of each signal measured using the ASSI metric defined earlier (Figure 3.11).

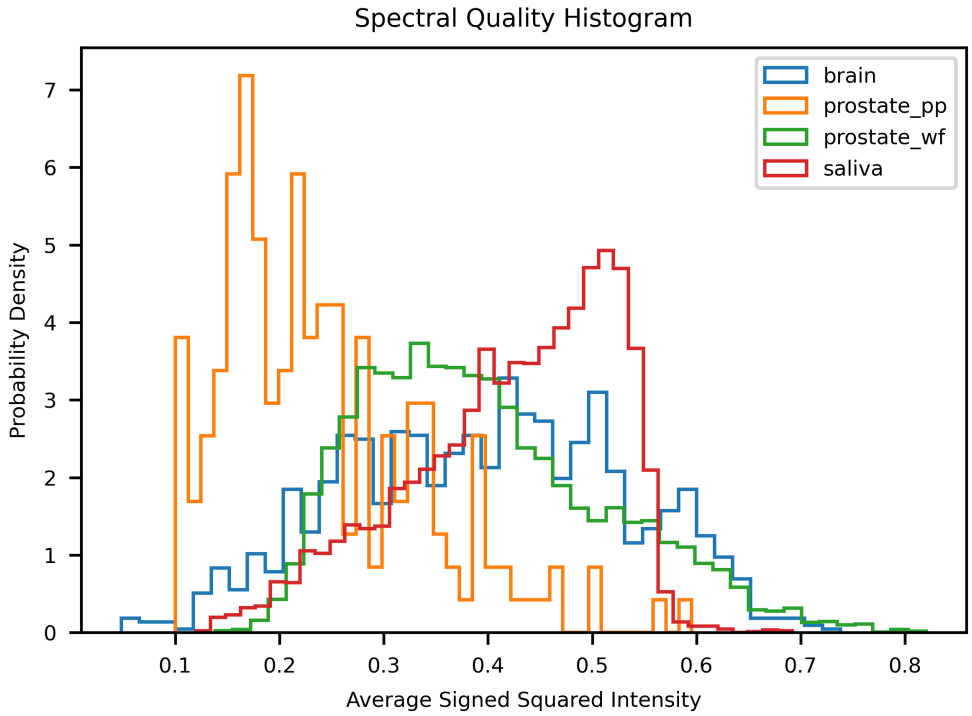


FIGURE 3.11 Spectral quality histogram as measured with the Average Signed Squared Intensity (ASSI) for each dataset. The ASSI is a spectral quality metric that is bound between -1 and 1. Using the ASSI, the overall quality of the spectrum from all datasets are compared. The `Prostate_PP` (e.g. prostate point-probe) dataset had the worst 'quality' spectrum (worse SNR and RBR of the 4 datasets). The Saliva dataset had the 'best' spectrum average signals. These four datasets covered a range of spectral quality representative of the biological sample landscape.

The raw data (first row of Figure 3.12) shows that the spectra of each dataset covered nearly the entire dynamic range of the acquisition instruments. This is frequent with biological samples and is explained by the large variability of fluorescence signal strength, absorption

and scattering. Nevertheless, Raman pre-processing using the workflow presented in this work and the BubbleFill algorithm for the removal of intrinsic fluorescence (step 3 of Figure 3.12) resulted in averaged spectra with small deviations. Many common Raman active bands can be identified and are common across all datasets, including vibrational modes typically found in proteins and lipids.

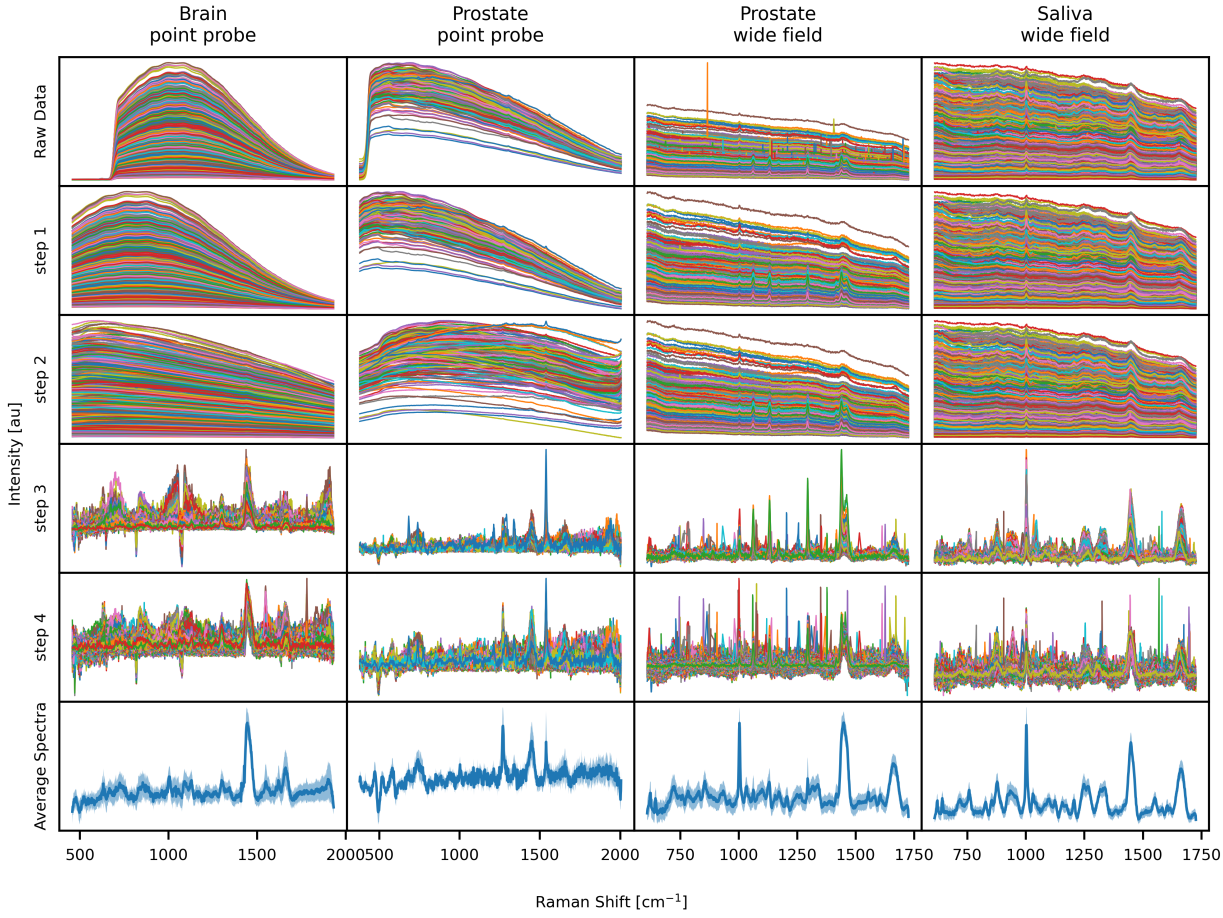


FIGURE 3.12 Raw Data : Measured raw spectra from instruments with accumulations combined. Step 1 : Truncation (for datasets 1 and 2) and cosmic ray removal. Step 2 : Background removal (for datasets 1 and 2) and calibration of x and y axis. Step 3 : Baseline removal with BubbleFill. Step 4 : Standard Normal Variate (SNV) normalization. Average Spectra : average spectra computed from step 4 results with ± standard deviation zone represented as shadow.

Finally, spectra of all datasets were clustered in groups of high, average and low quality based on the ASSI metric (Figure 3.13). As the quality increased, the standard deviation became smaller, converging toward the average spectrum. This behaviour was observed for all datasets. Furthermore, some specific Raman bands (peak at 1300 cm^{-1} for the prostate wide

field dataset) gained in intensity while artifacts (400 to 600 cm^{-1} for the prostate point probe dataset) disappeared. These results indicate that the ASSI metric has the potential to be used in machine learning applications to discard low-quality spectra based on a threshold or as a general signal quality metric during acquisition to facilitate troubleshooting of instrument and software.

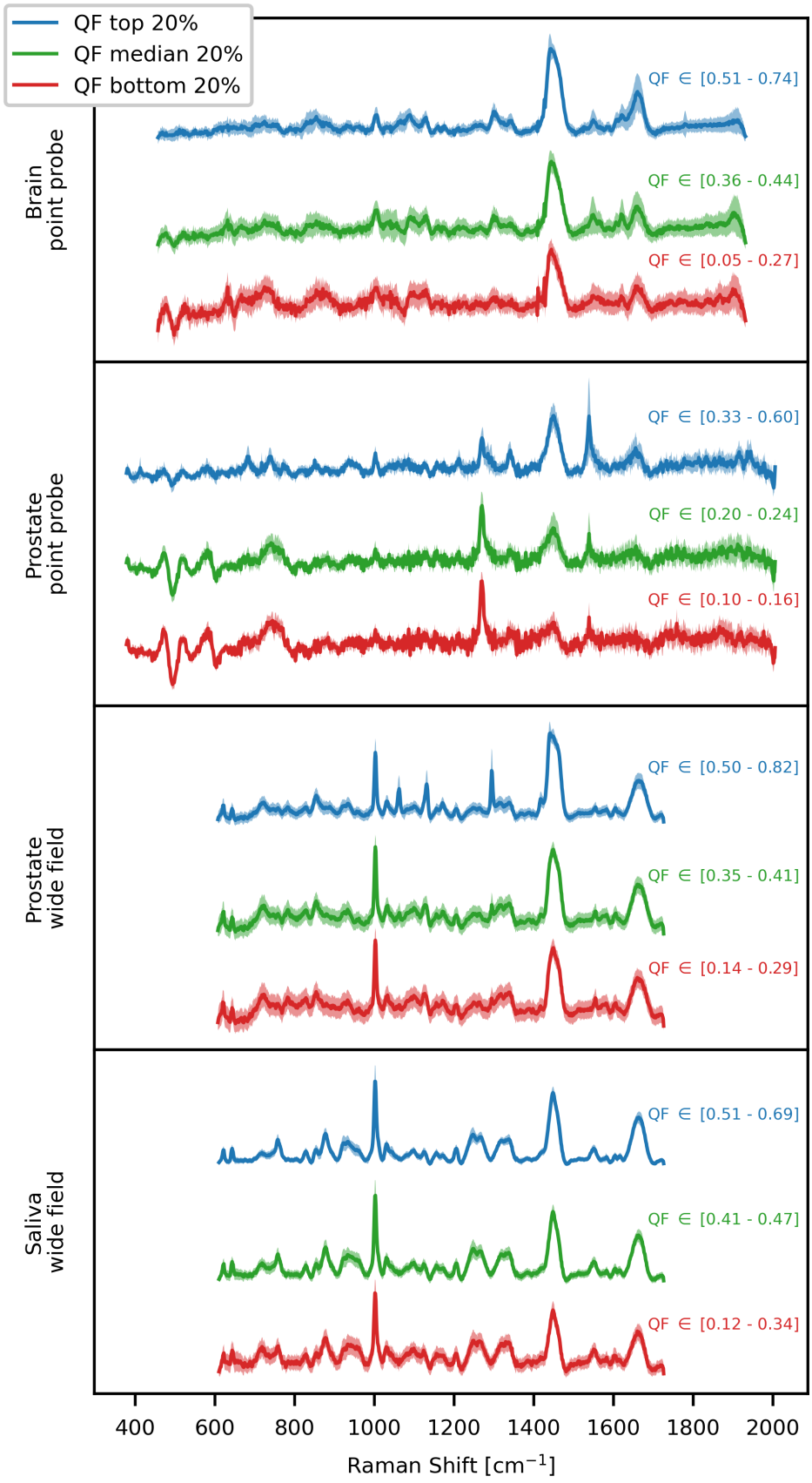


FIGURE 3.13 Average pre-processed spectra of each experimental tested datasets clustered in high (top 20 % ASSI), average (middle 20 % ASSI) and low (bottom 20 % ASSI) quality.

3.5 Conclusion

In conclusion, we developed and released a python package under the MIT license that implements the necessary tools for Raman spectra pre-processing. Most notably, the ORPL package includes the novel BubbleFill algorithm intended for the removal of autofluorescence baselines. We validated BubbleFill using a combination of numerical benchmarks based on synthetic spectra and real-world experimental data from previous studies. Comparative benchmark results revealed that BubbleFill performed better, especially for the removal of an aluminium baselines, which is of critical importance in many of our group's studies, or as well as other commonly used algorithms. Although these results might not generalize to every possible baseline shapes, the quantitative comparison methodology presented in this work can be extended to include a larger variety of fluorescence and Raman responses. This makes it ideal for selecting and tuning baseline removal algorithms for specific applications while limiting user biases. Finally, we combined the Raman acquisitions of previous studies in a single dataset which is, to our knowledge, the largest and most varied Raman dataset compiled for the purpose of testing and validating pre-processing. This data was used to validate the ORPL package on signals covering a wide range of signal-to-noise and signal-to-background ratios representative of the biological application landscape. In the future, additional modules will be added to ORPL to address other critical challenges such as spectral unmixing and peak analysis tools and chemometric analysis. It is our hope that this package be used as a stepping stone enabling a more open and uniformed pre-processing methodology across the Raman research and clinical spectroscopy scientific communities.

3.6 Acknowledgements

This work is supported by the Discovery Grant program from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Collaborative Health Research Program (CIHR and NSERC), Mitacs and the Canada First Research Excellence Fund (TransMedTech Institute). We would like to thank the company Reveal Surgical for giving access to part of an in-human brain cancer dataset.

3.7 Conflict of interest

We do not declare any conflict of interest concerning this study.

3.8 Code, Data and Materials Availability

The ORPL package is hosted on the python package index (PyPi) as an open sourced project distributed under the MIT license and the name ORPLIB (`https://pypi.org/project/orplib/`). The source code is available on GitHub at `https://github.com/mr-sheg/orpl`. The data and materials information that support the findings of this study are available from the corresponding author upon reasonable request.

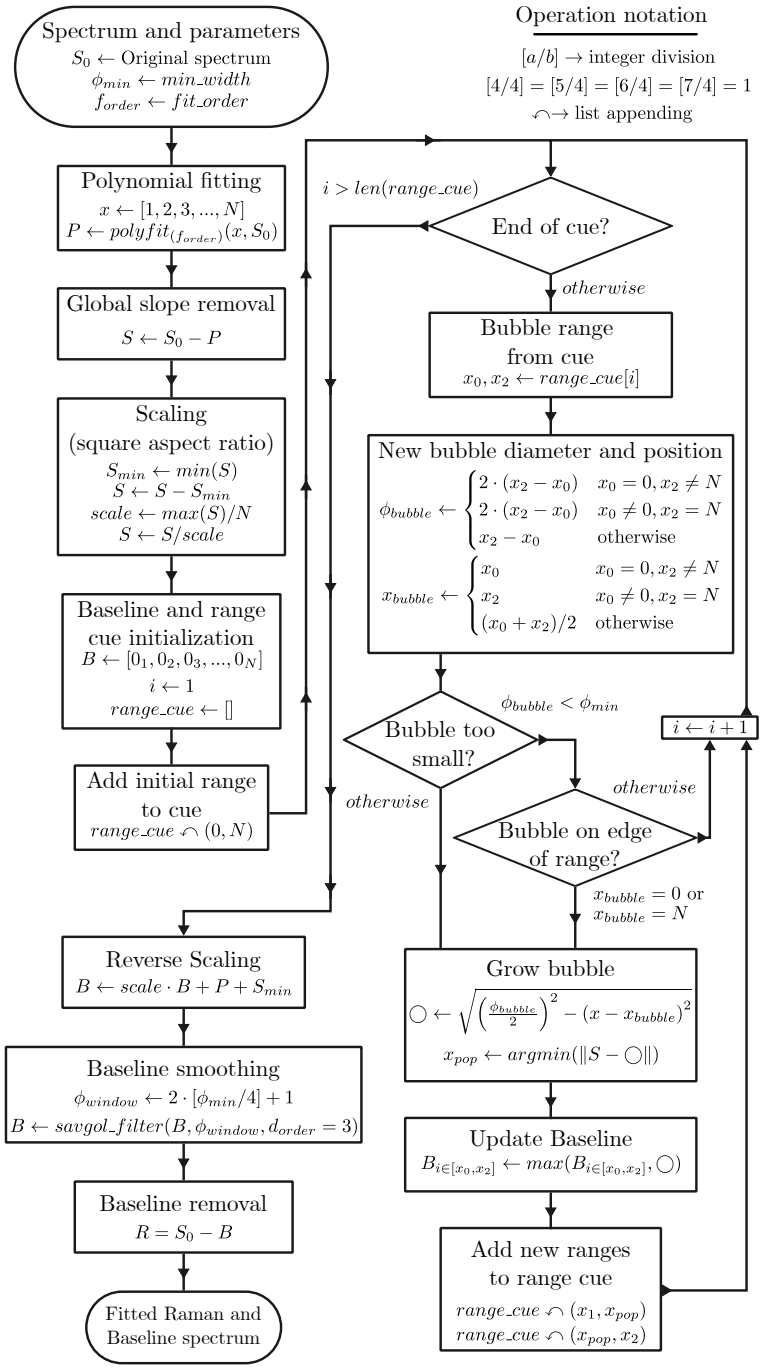


FIGURE 3.4 Flowchart of the BubbleFill baseline removal algorithm. The algorithm inputs are : a) A Raman signal (S_0) represented as a vector of length N, b) the minimum width of the bubbles allowed to grow (ϕ_{min}) and c) the order of the polynomial fit used to remove the global slope of the signal (f_{order}). First, the global slope of the signal (S_0) is removed using a polynomial fit and the signal intensity is normalized to obtain a square aspect ratio ($\min(S_0) = 0$, $\max(S_0) = N$). Then, bubbles of increasingly smaller size are grown underneath the signal and the baseline fit (B) is iteratively updated. Once the bubble growth loop is completed, the square aspect ratio normalization is reversed and the baseline fit is smoothed using a Savitzky-Golay filter. The output Raman signal is computed as the subtraction of the baseline fit to the input signal.

CHAPTER 4 THE OPEN LUMED RAMAN PLATFORM

4.1 Scientific and Technical Context

For the past decade, the Lumedlab (formerly the Laboratory of Radiological Optics) research group has been at the forefront of integrating Raman spectroscopy into clinical settings. Numerous studies conducted by the group have demonstrated the potential of Raman technologies to enhance diagnostic capabilities, particularly in real-time surgical environments [9, 55, 19, 137, 51, 48, 56, 5].

However, despite these advancements, the integration of Raman systems into clinical workflows has encountered significant challenges. A critical issue arises from the need for rapid setup and preparation times, typically within a few minutes, along with the requirement for repeatable, safe instrument configurations that minimize user input. While these features are essential for seamless operation in dynamic settings, such as operating rooms, they often result in compromised measurement accuracy and reliability. Additionally, ensuring data consistency across different systems has proven to be a persistent challenge[135]. Variations in instrumentation and environmental conditions can lead to discrepancies in the acquired spectra, complicating the interpretation and comparison of data. To better understand and monitor these variations, spectrum quality metrics were developed, which serve as tools for evaluating the reliability of acquired data. This work has underscored the importance of maintaining consistent data quality, particularly when deploying Raman systems in diverse clinical and research environments[137].

In practice, the strength of Raman signals can vary substantially between samples or across different interrogation sites, which complicates manual adjustments to acquisition parameters and increases the likelihood of suboptimal measurements. These limitations have hindered the consistent performance of Raman systems in clinical applications, particularly in fast-paced environments where precision and efficiency are paramount.

The work presented in this chapter addresses these challenges by introducing the Lumed Raman Platform. This platform integrates the Lumed Raman System (LRS), an open-source Raman acquisition system specifically designed for clinical diagnostics and research, with the Open Raman Acquisition Software, a Python-based software platform for controlling Raman acquisition systems. The combination of these technologies represents a significant advancement in the reliable deployment of Raman spectroscopy in real-world surgical settings, enhancing both measurement accuracy and operational ease.

4.2 Laser Safety Considerations

As with any light-based technology intended for clinical use, the development of the Lumed Raman Platform required careful consideration of laser safety. This is particularly crucial in neurosurgical applications, where the system may be used intraoperatively on exposed brain tissue. In such contexts, even minimal photothermal or photochemical damage can have serious consequences. Ensuring safe operation is therefore essential—not only for regulatory compliance, but also to protect patient well-being. A key aspect of this is understanding and applying the concept of maximum permissible exposure (MPE), which defines the safe limits for laser irradiation. While the system may occasionally operate above MPE thresholds, particularly in cases where diagnostic benefit outweighs the associated risk, it is critical that clinicians understand how such exposures influence the likelihood of tissue damage. By providing a clear framework for assessing and managing these risks, the Lumed Raman Platform supports informed decision-making in the operating room, enabling safe and effective integration of Raman spectroscopy into surgical workflows.

4.2.1 Maximum Permissible Exposure

The maximum permissible exposure (MPE) is defined by the American National Standards Institute (ANSI) as "The level of laser radiation to which an unprotected person may be exposed without adverse biological changes in the eye or skin."[138] It corresponds to a hard limit for the level of laser radiation to which an individual may be subjected under normal circumstances. However, this limit can be exceeded for patients when the exposure is intended or acceptable by health care personnel for therapeutic or diagnostic purposes whether there is a risk or certainty of tissue damage[139].

The MPE is calculated differently based on the laser wavelength and duration of the exposure. For a continuous exposure of duration t longer than 100 ns (10^{-7} s) from a laser of wavelengths 400 nm to 1400 nm, the MPE for skin is

$$\mathrm{MPE}_{skin} = \begin{cases} 1.1 C_A t^{0.25} & \mathrm{J/cm^2} \quad t \in 10^{-7} \text{ to } 10\mathrm{s} \\ 0.2 C_A & \mathrm{W/cm^2} \qquad t > 10\mathrm{s} \end{cases} \tag{4.1}$$

where C_A is a wavelength dependent correction factor given as

$$C_A = \begin{cases} 1.0 & \lambda \in 400 \text{ to } 700\text{nm} \\ 10^{0.002(\lambda - 700)} & \lambda \in 700 \text{ to } 1050\text{nm} \\ 5.0 & \lambda \in 1050 \text{ to } 1400\text{nm} \end{cases} \tag{4.2}$$

For a continuous exposure of duration t longer than 5 µs (5×10^{-6} s) from a laser of wavelengths 700 nm to 1050 nm, the MPE for the eye is

$$\text{MPE}_{eye} = \begin{cases} 1.8 C_A C_E t^{0.75} \times 10^{-3} & \text{J/cm}^2 \quad 5 \times 10^{-6} < t \leq T_2 \\ 1.8 C_A C_E T_2^{-0.25} \times 10^{-3} & \text{W/cm}^2 \quad t > T_2 \end{cases} \tag{4.3}$$

C_A is the same correction factor as for the skin and C_E and T_2 are the correction factors for a circular extended source

$$C_E = \begin{cases} 1.0 & \alpha > \alpha_{min} \\ \alpha / \alpha_{min} & \alpha_{min} \leq \alpha > \alpha_{max} \\ \alpha^2 / (\alpha_{max} \alpha_{min}) & \alpha_{max} \leq \alpha \end{cases} \tag{4.4}$$

and

$$T_2 = 10 \times 10^{(\alpha - 1.5)/98.5} \quad \text{s} \tag{4.5}$$

α, α_{min} and α_{max} are subtense angles given in mrad. The angular subtense is the angle that corresponds to the apparent visual angle as calculated from a source's size and distance from the eye. For a circular source of diameter d at distance l, this angle is

$$\alpha = 2 \cdot \arctan(\frac{1}{2} \cdot \frac{d}{l}) \simeq \frac{d}{l} \tag{4.6}$$

α_{max} depends on the duration of the pulse as

$$\alpha_{max} = \begin{cases} 5 & \text{mrad} \quad t > 625\text{µs} \\ 200 t^{0.5} & \text{mrad} \quad 625\text{µs} > t \leq 0.25\text{s} \\ 100 & \text{mrad} \quad 0.25\text{s} \leq t \end{cases} \tag{4.7}$$

with $\alpha_{min} := 1.5\text{mrad}$.

An extended laser source can effectively be considered a point source if it has a subtense angle $\alpha > \alpha_{min}$. As an example, this corresponds to any source which apparent visual angle is equivalent (or smaller) to that of a circular source of diameter $d = 1.5$ mm viewed at a distance $l = 1$ m. Considering all laser radiation sources as point source is an approximation that is conservative (i.e. the MPE for a point source is lower than that of an equivalent extended source) and appropriate for the work presented in the rest of this document. When making this approximation, the MPE for the eye simplifies to

$$\mathrm{MPE}_{eye} = \begin{cases} 1.8C_A t^{0.75} \times 10^{-3} & \mathrm{J/cm^2} \quad 5 \times 10^{-6} > t \leq 10\mathrm{sec} \\ C_A \times 10^{-3} & \mathrm{W/cm^2} \quad t > T_2 \end{cases} \tag{4.8}$$

4.2.2 Maximum Exposure Duration

Currently, the MPE has been rigorously established only for two organs : the eye and the skin. MPE limits are not specifically defined for other tissues in current guidelines. In situations where laser exposure occurs near the eye, additional precautions are required to mitigate risks. This falls outside the scope of the work presented here. Given that light radiation exposure during Raman spectroscopy generally targets tissues other than the eye, the MPE for skin is used as the chosen limit. This provides a safe basis for analysis of laser exposure risk factors, but does not exclude that additional considerations may be required with laser exposure to other tissues.

The limiting aperture diameter used for measuring laser irradiance and assessing the risk factor for skin exposure is $D = 3.5$mm, regardless of the size of the region exposed to light [138]. If the region of skin tissue exposed to radiation exceeds this limiting aperture, the average power surface density should be calculated based on an equivalent 3.5 mm spot size. In other words, if the spot size is larger than 3.5 mm, the total power should be averaged over an area corresponding to a 3.5 mm diameter to compare with the MPE. This ensures that the exposure is evaluated using a consistent reference, regardless of the actual size of the irradiated area.

Since all probe instruments used in the LRS have an output window diameter larger than $D = 3.5$mm, the effective illumination area considered for comparison with the MPE is $A = 9.62 \times 10^{-6}\mathrm{m}^2$. Therefore, an arbitrary light exposure (E) can be calculated as :

$$E = \frac{P \times t}{A} \mathrm{J/m^2} \tag{4.9}$$

where P is the output power of light at the probe tip, and t is the duration for a continuous

exposure. Equations 4.1 and 4.9 can be combined to determine the relationship between power and duration of laser excitations for different MPE multiplication factors. These multiplication factors — denoted as scalar values of 1x, 2x, or 5x — scale the MPE limit. Each multiplication factor is represented by a distinct curve in Figure 4.1, illustrating how varying exposure thresholds impact the overall risk factors for tissue damage.

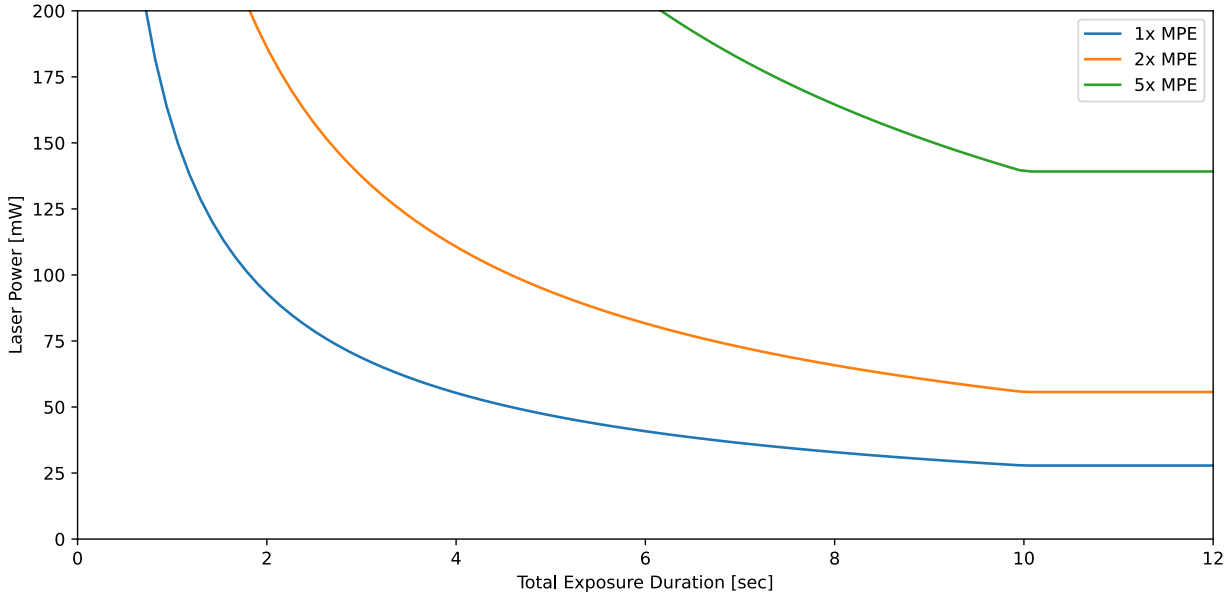


FIGURE 4.1 Relationship between laser exposure power and duration for MPE multiplication factors 1x, 2x, and 5x.

4.3 The Lumed Raman System

The Lumed Raman System (Figure 4.2) comprises an EMVision LLC spectrometer and an Andor Newton CCD low noise high-efficiency camera for spectrum measurement. The Raman excitation source is an IPS 785 nm fibered laser (Multi-mode digital U-Type module) and a Thorlabs PM100USB powermeter is used for optical power measurement and calibration. An optional Arduino control board can be integrated to enhance security features such as emergency shut-off of the laser. The system further supports various fibre optic probe instruments, which can be connected to the spectrometer via an industry-standard MTP keyed connector with vertically aligned optical fibres and to the laser via an SC/PC connector, ensuring seamless integration and operation. A detailed guide that outlines the assembly procedure of the system is available in Appendix A.

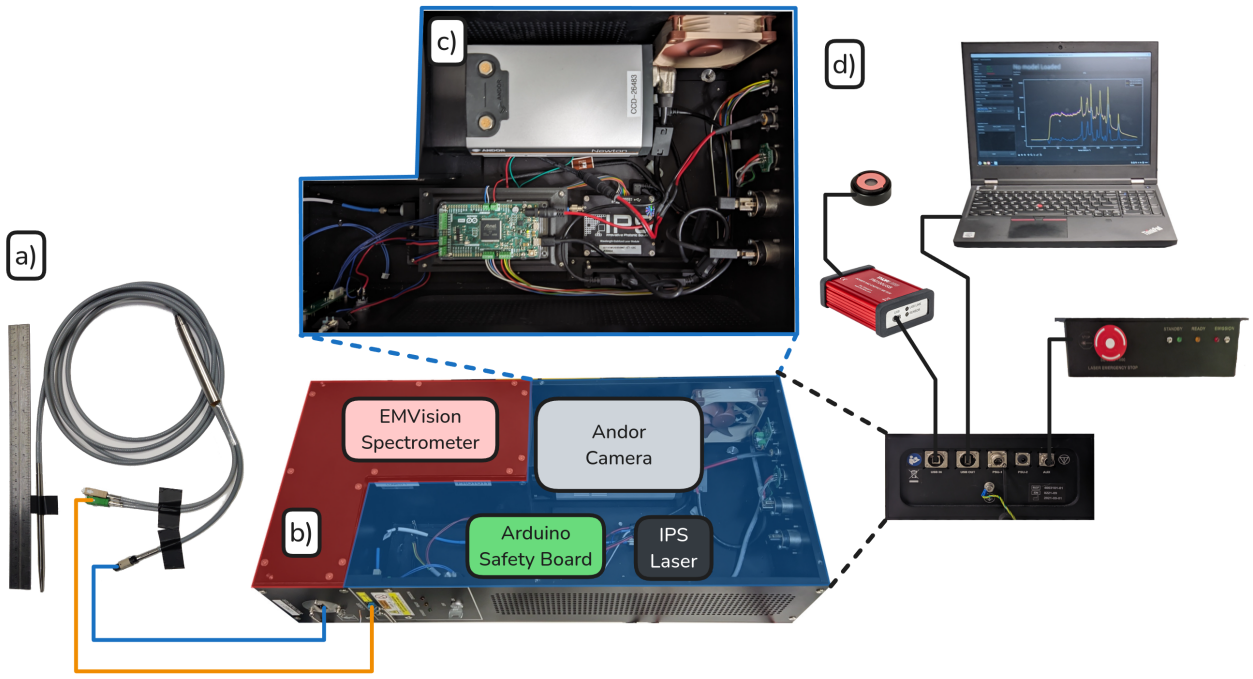


FIGURE 4.2 Lumed Sentry system component overview. (a) EM Vision Fibre optics based probe. (b) EM Vision HT spectrometer. (c) Internal electronics (Andor Newton CCD camera, IPS laser, Arduino board, usb hub, ...). (d) External connections to a control computer, the auxiliary box and a Thorlabs powermeter.

4.3.1 The EMVision Probe

The optical probes used with the Lumed Raman System are manufactured by EMVision LLC and have been optimized for measurement of Raman spectra in a back scattering geometry. The probe design, engineering and manufacturing details are protected under several patents[140, 141, 142, 4] and involved years of development and optimization. One variation of the EMVision probe is illustrated in Figure 4.3. Excitation light from a laser is carried to the tip of the instrument through an excitation fibre (110). The light passes through an emission fibre filter (109) that filters-out Raman contribution from the fibre and cleans up the laser-line. Then, the excitation light reaches the sample after passing through the probe tip window (101). Backscattered Raman light from the sample passes back through the probe window (101), a lens filter (102), a lens (104) and a collection fibre filter (106) before being guided back the length of the probe by the collection fibres (107).

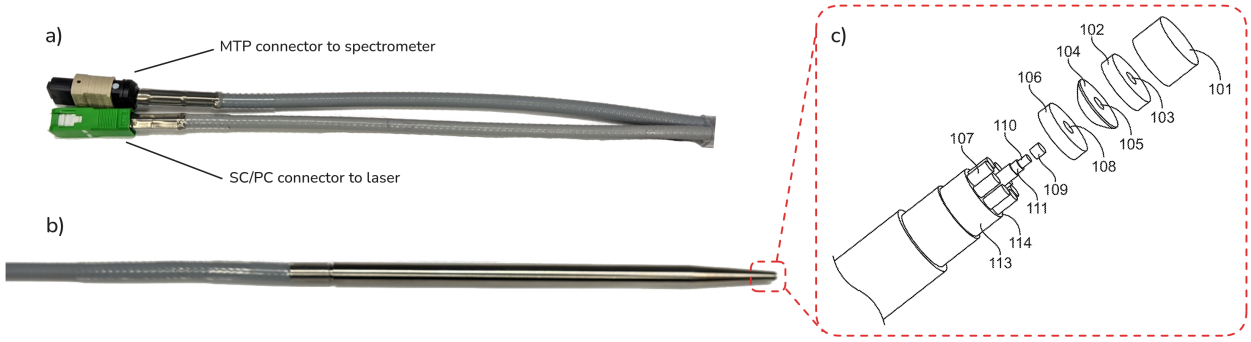


FIGURE 4.3 EMVision *Laboratory Probe*. (a) Fibre connector to the Lumed Raman System laser (SC/PC) and spectrometer (MTP). (b) Probe main body. (c)Schematic diagram of the optical component assembly at the probe tip (reproduced from [4])

Various probe design are possible and have been used by the Lumed lab in different context and for different studies over the years. The probe design is determined by the practical challenges specific to the clinical context, and often limited by the surgical field and the difficulty of reaching sites or tissue of interest with the instrument. The probe most commonly used by the Lumed lab is the *Laboratory Probe* (Figure 4.4.a). It features a 2.1 mm probe tip, has a single laser emitting fibre (usually 300 µm) and 7 to 9 collection fibres (usually 300 µm). The *needle probe* (Figure 4.4.b) has a rigid section 20 cm in length, an outer diameter of 1.22 cm and used a 300 µm excitation fiber and 10 x 100 µm collection fibres. A variation of the *needle probe* was used to take measurements during prostate brachytherapy procedures through the needle insertion guiding template [51, 52]. Another custom-made probe (Figure 4.4.c) was more recently used during transsphenoidal endonasal endoscopic resection surgery to take measurement of pituitary adenomas and other tumours [5]. In this case, the extreme challenge posed by the difficulty of reaching the tissue sites of interest required a probe with a bifurcation near the distal end and a long and narrow tip. Lastly, flexible *endoscopic probes* can be manufactured by encapsulating the fibres in a rubber jacket instead of a stainless-steel rod (Figure 4.4.d).

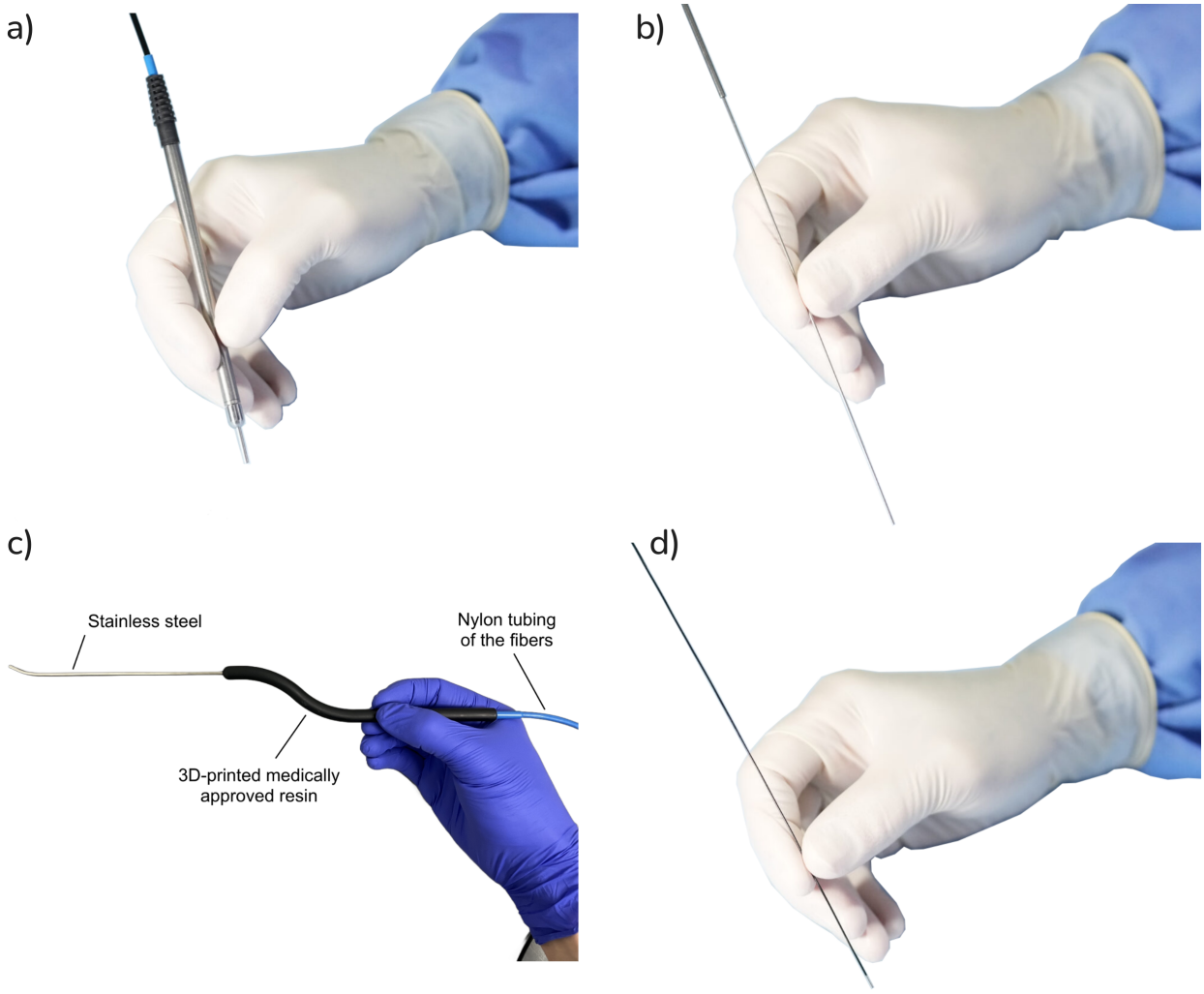


FIGURE 4.4 Pictures of the different EMVision probes that can be used with the Lumed Raman System. (a) is the *Laboratory Probe*. (b) is the *needle probe*. (c) is a custom probe designed for measurements during transsphenoidal endonasal endoscopic resection surgery [5]. (d) is the *endoscopic probes*.

4.3.2 The EMVision Spectrometer

The EMVision LLC model HT spectrometer is based on a simple yet sophisticated design registered under several patents [143, 144, 145, 6]. An example configuration is illustrated in Figure 4.5. It comprises a main housing (101), a slit (102), a collimating lens (104), a filter (106), a transmission Volume Phase Holographic Grating (108) and a focus lens (110). Light collected from the probe connected to the entry port first passes through a narrow vertical slit and is collimated to the optical axis by the collimating lens. The filter blocks residual laser light, while the collected Raman light ($\lambda > 785$ nm) can pass through. The

grating introduces wavelength dependent angular separation of the light along the horizontal plane (x-axis of the camera). Finally, the light is focused on the camera array sensor by the focusing lens. Specifications reported by EMVision LLC for the spectrometer package used in the Lumed Raman Systems (100 µm slit equipped with an Andor Newton camera) are presented in Table 4.3.

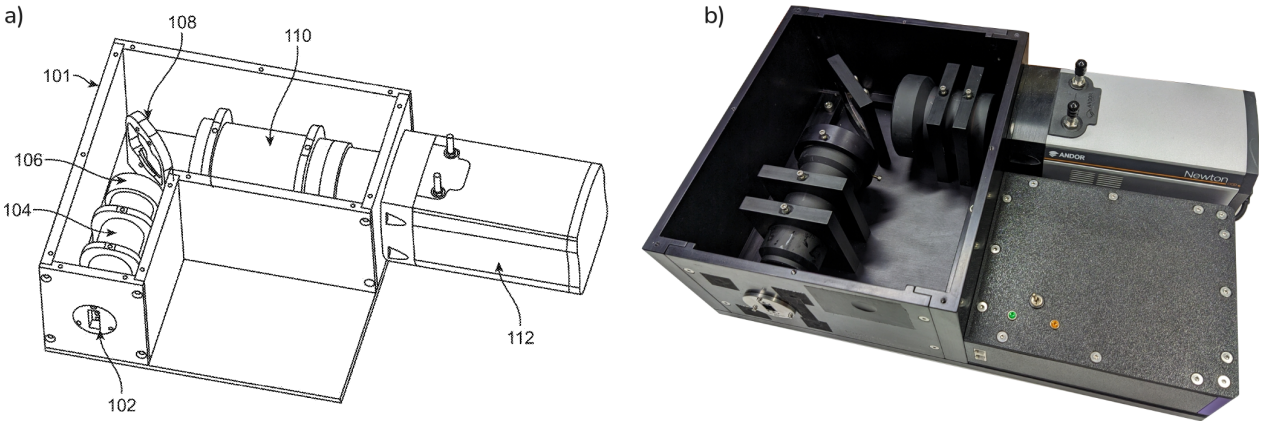


FIGURE 4.5 EMVision LLC model HT spectrometer. (a) Schematic diagram illustrating the main housing (101), entrance slit (102), collimating lens (104), laser-rejection filter (106), transmission Volume Phase Holographic Grating (108), and a focusing lens (110) (reproduced from [6]). (b) Photograph of the assembled spectrometer unit as implemented in the Lumed Raman Systems

TABLE 4.1 Technical specifications of the EMVision LLC model HT spectrometer when equipped with an Andor Newton CCD camera

Technical Specification	
Raman Excitation Wavelength	785 nm
Spectral Range	350 cm^{-1} to 2100 cm^{-1} (807 nm to 940 nm)
Slit Size / Resolution	100 µm slit, $\geq$8.7 cm^{-1} spectral resolution
Laser Blocking Filter	Optical Density >6 at 785 nm
Grating	Volume Phase Holographic Transmission
Probe Input Connector	Keyed snap-in MTP linear connector (up to 9 linear 300 micron core fibres)

When used in tandem with the EMVision probes, the images formed on the camera sensor array show several horizontal bright lines. Each of the bright lines corresponds to a single collection fibre of the probe, and the separation between the lines corresponds to the spacing of the fibres on the probe connector. Accordingly, the camera sensor's horizontal axis maps to the wavelength axis of the spectrometer and its vertical axis maps to the vertical spatial axis of the probe connector (Figure 4.6).

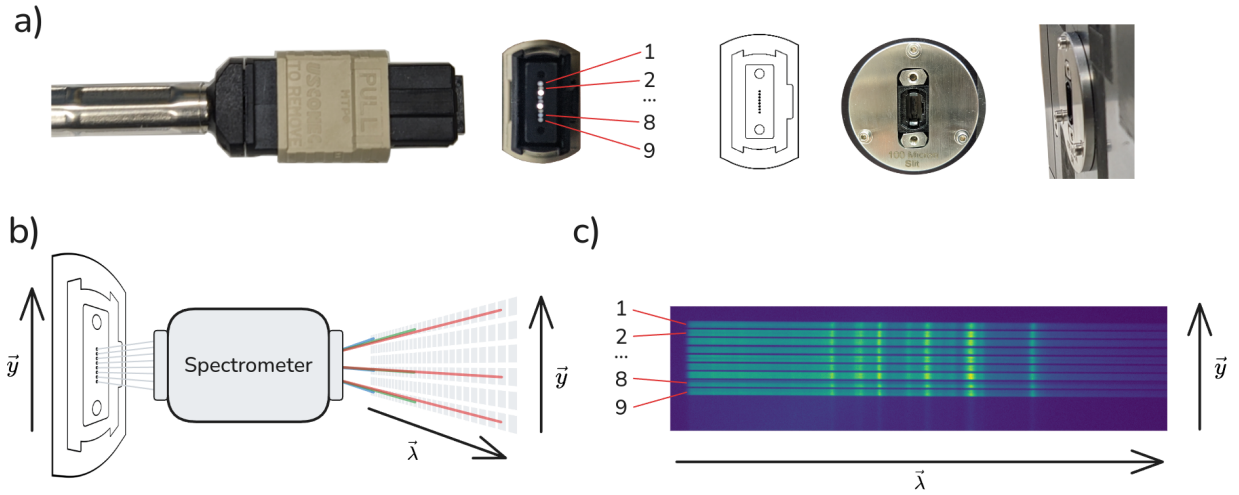


FIGURE 4.6 (a) Optical connection between the probe and spectrometer, showing the probe's MTP connector (left) and spectrometer input port mounting plate (right). (b) Schematic of vertical fibre alignment and spectral dispersion, illustrating the mapping of wavelength (x-axis) and fibre separation (y-axis) information onto the camera sensor. (c) Example sensor image, showing nine horizontal spectral lines, each corresponding to a separate collection fiber.

4.3.3 The Andor Newton Camera

The camera used in the Lumed Raman System is an Andor Newton CCD camera model DU920P_BR-DD that is ideally suited for low signal applications in the visible-NIR wavelength range. It has a peak quantum efficiency of 95 % while maintaining minimal noise level and fast acquisition speed. Summary specifications of the camera are presented in Table X and the quantum efficiency curves for different Andor camera models are presented in Figure 4.7.

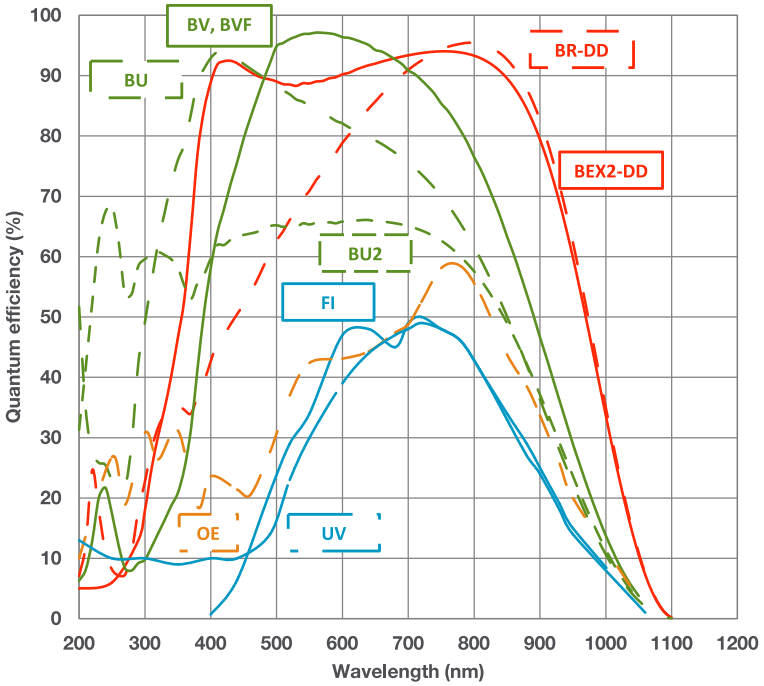


FIGURE 4.7 Quantum efficiency curves of Andor Camera models.

The camera can be operated using a number of different binning patterns or *readout modes* (Figure 4.8) and a number of acquisition sequences or *acquisition modes* (Figure 4.9). The readout mode chosen for an acquisition specifies how the array of pixels is read. For instance, in Full Vertical Binning (FVB), the entire chip is used as if it was a single linear image sensor. All the charges from a single column of pixel are binned into the shift register and counted only once. This mode significantly reduces read noise in the measured signal, which would result in a 1024 data points vector (for the Andor Newton). Other read modes provide different levels of control over binning of the chip (Table X). In addition, horizontal binning can be used when configuring the read mode of the camera to combine pixels horizontally. This can provide an x-fold increase in signal at the cost of an equivalent loss in spatial resolution — or spectral resolution when using the camera with a spectrometer.

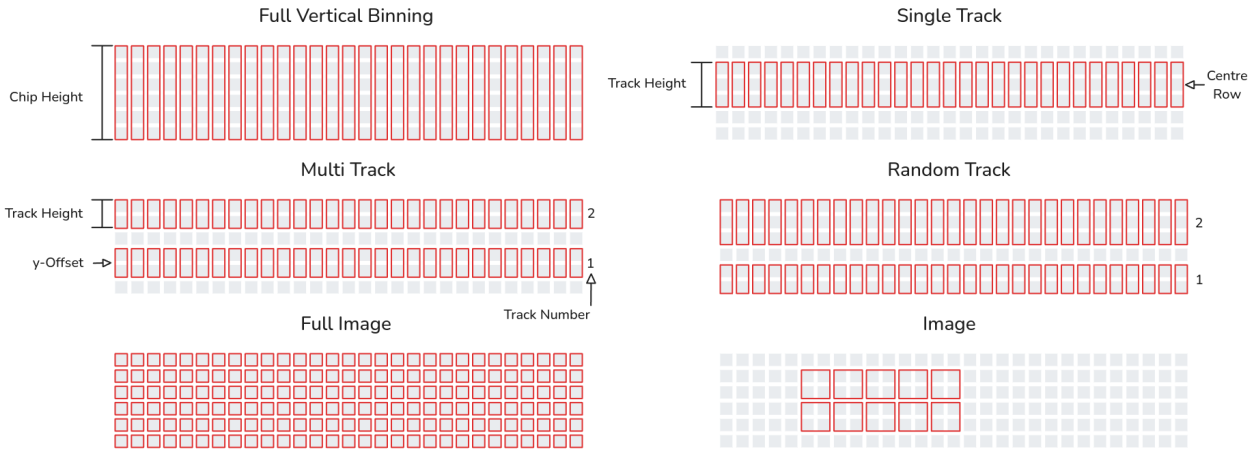


FIGURE 4.8 Illustration of available readout modes for the Andor Newton camera.

TABLE 4.2 Overview of readout modes available for the Andor Newton camera and their associated control parameters.

Read Mode	**Control Parameter**
Full Vertical Binning	None
Single Track	Track Height and Centre Row
Multi-Track	number of tracks, track height (same for all tracks), and track offset (from the bottom track)
Random Track	a list of track start and track end index(`[track_1_start, track_1_end, track_2_start, track_2_end, ...]`)
Full Image	None
Image	`(x,y)` coordinates of the bottom-left and top-right corners and the number of super pixels to use in the x and y dimensions (must be divisible in an integer number of pixel)

The acquisition mode provides control over the acquisition sequence. The simplest of which is a single scan, where a single readout of the CCD is performed. The only parameter to configure in this mode is the exposure time of the time-integration window. Although simple, this mode is ill-suited for repeated measurements in quick succession, as the data from the scan needs to be sent to the PC after each readout. An accumulated scan is equivalent to taking multiple single scans and adding the results together manually. However, the accumulated scan provides fine control over the time delay between the consecutive scans. The best acquisition mode for the quick, repeated accumulations used in Raman spectroscopy is the Kinetic Series. In this mode, an acquisition consists of a series of N identical readouts. The data of each scan is recorded in the camera's memory and sent to the PC at the end of the acquisition. For instance, a kinetic series of 3 cycles of 0.3 s exposure time with a kinetic

cycle time of 3 s would take 9 s to complete and return 3 spectra, each with a 0.3 s exposure duration to the PC whereas an equivalent accumulated scan would return a single spectrum where the cumulative data from each scan is summed. Control parameters available for the different acquisition modes supported by Andor cameras are summarized in Table X. In practice, the Lumed Raman System uses Kinetic series accumulation with the shortest possible kinetic cycle time for the measurement of spectra.

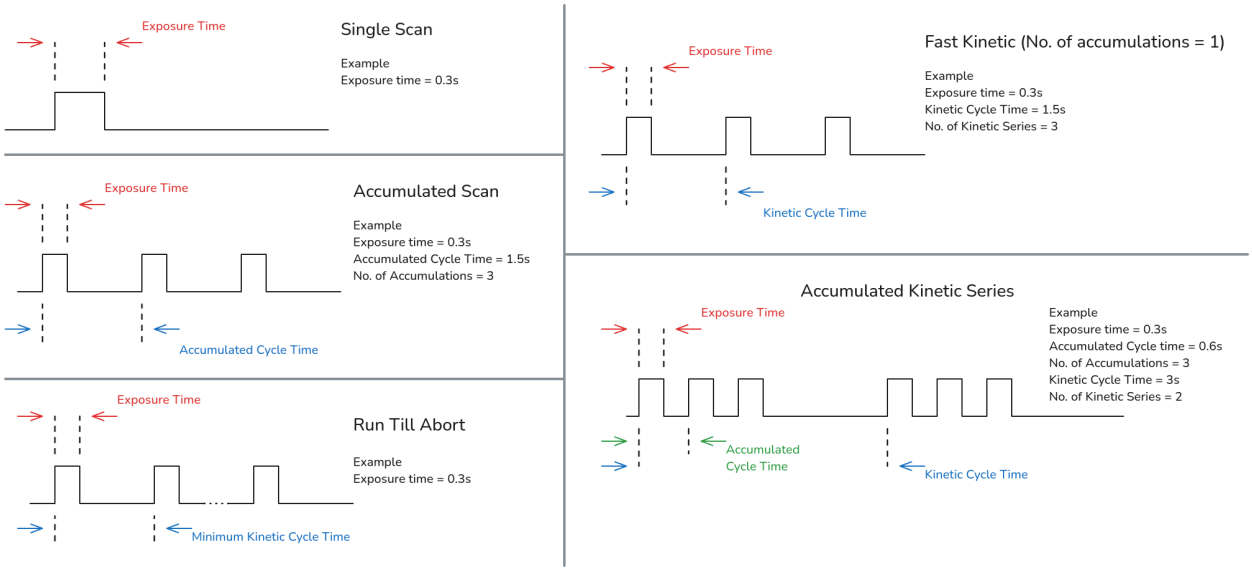


FIGURE 4.9 Overview of the camera's acquisition modes, detailing the time sequence and control parameters.

TABLE 4.3

Acquisition Mode	**Exposure Time**	**Accumulated Cycle Time**	**No. Of Accumulations**	**Kinetic Cycle Time**	**No. Of Kinetic Series**
Single Scan	✓				
Accumulated Scan	✓	✓	✓		
Kinetic Series	✓	✓	✓	✓	✓
Run Till Abort	✓			✓	

Software communication with the camera can be made with Andor's own proprietary software, Andor Solis or with a custom application through the Andor's Software Development Kit (SDK) and the Andor instrument drivers.

4.3.4 The IPS Laser

The laser light source used in the Lumed Raman Systems is an Innovative Photonics Solution (IPS) 785 nm multimode digital U-type module. This laser model is available in three versions : `I0785MU0350MF-USB`, `I0785MU0500MF-USB` or `I0785MU0800MF-USB` which are rated for a maximum output power of 350 mW, 500 mW and 800 mW respectively. Optical specifications of the IPS laser are presented in Table 4.4 and the laser electrical pinout is illustrated in Figure 4.10 and detailed in Table 4.5.

TABLE 4.4 Optical specifications of the IPS multimode digital U-type module laser

Optical Specification	
Wavelength Tolerance	±0.5 nm
Spectral Linewidth	0.8 nm
Wavelength Stability Range	15 °C to45 °C
Output Power Stability	<1 % at constant case temperature Modulation Rate & Digital Modulation : 25 kHz External Modulation : CW to 10 kHz at 50 % duty cycle or CW to 1 kHz at 10 % –100 % duty cycle.
Warm-Up Time	10 s from cold start or1.5 s from warm start

TABLE 4.5 Electrical pinout of the IPS multimode digital U-type module laser

Pin number	Pin Label	Description
1	`V+`	Power Supply : 5 V to 12 V (12 V recommended)
2	`VBIAS ENABLE` (Set Enable)	For analog control : Enable external laser power control through pin 8 (`LD VBIAS`)
3	`SCL/TX`	I2C : SCL standard (RS232 : TX ; board by request)
4	`SDA/RX`	I2C : SDA standard (RS232 : RX ; board by request)
5	`GND`	Ground
6	`V+`	See Pin 1
7	`Enable`	Laser Enable : `HIGH` (3.3 V to `V+` supply voltage) = Enable `LOW` (`GND`) = Disable
8	`LD VBIAS` (LD Set)	Laser power set point – Enables analog external control of laser drive current
9	`PD+`	For analog readout. Connect voltmeter to `PD+` (pin 9) and `GND` for photo diode V output (0 V to 3.3 V)
10	`GND`	Ground

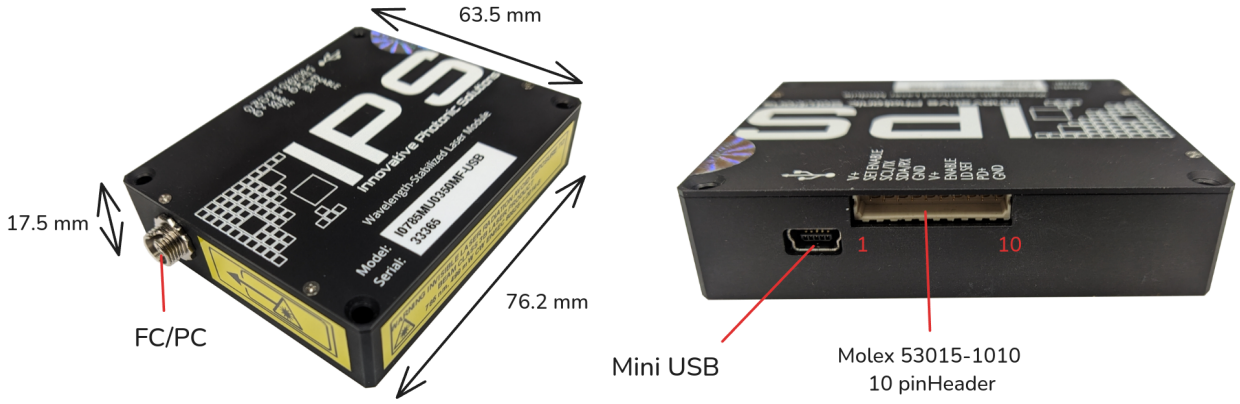


FIGURE 4.10 Innovative Photonics Solution (IPS) 785 nm multimode digital U-type module laser. Physical dimensions and fibre connector on the left and USB and electrical pinout MOLEX connector on the right.

Power must be supplied to both `V+` pins (pin 1 and 6) and ground must be applied to both `GND` pins (pin 5 and 10) for the unit to be turned ON. Pin 7 must be set to `HIGH` for the laser to be enabled. Once powered and enabled, the laser can be controlled in analog mode by modulating the voltage on pin 2 while pin 8 is set to `HIGH` or in digital mode via serial commands through the mini-USB port or with I2C communication through pin 3 and 4. IPS lasers in the Lumed Raman System are controlled in digital mode via serial communication through the mini-USB port of the unit. The command syntax follows the Standard Commands for Programmable Instruments (SCPI v1999.0) and is detailed in the IPS user guide for the Digital U-Type units. A few important commands are described in Table 4.6.

TABLE 4.6 Command reference summary for the Innovative Photonics Solution (IPS) 785 nm multimode digital U-type module laser

Command	Description
`"*IDN?"`	Reports the device identification string. Will read back : `IPS, HPU, laser serial number, factory measured wavelength, FW-revision`
`"Board:Current?"`	Reports the measured current draw in mA
`"Board:Temperature?"`	Reports the module case temperature in °C
`"Laser:Current <current>"`	Sets laser operating current set point to `<current>` mA
`"Laser:Current?"`	Reports measured laser operating current in mA
`"Laser:Enable <enable state>"`	Enables/Disables laser operation (`<enable state>` : `1` to enable or lstinline|0| to disable)
`"Laser:Enable?"`	Reports laser enable state
`"Laser:Temperature?"`	Reports the measured laser/TEC temperature (°C)

4.3.5 The Laser Safety Control Board

The Laser Safety Control Board of the Lumed Raman System is the component responsible for enabling and disabling the laser when the appropriate conditions are met and reporting the laser state to the operator. It functions as a stand-alone sub-system and does not rely on communication with the computer that operates the rest of the main system. Nevertheless, it can still report information regarding its state via serial communication for debugging and troubleshooting purposes. This ensures that software or computer issues can never impede the ability of the control board to quickly shut off laser emission in case of emergency, and for the operator to assess the state of the laser by reading the LED display on the front panel of the unit. Figure 4.11 illustrates a schematic of the electronic connections between the control board and the indicator elements (LEDs), control elements (switches and buttons) and the laser module.

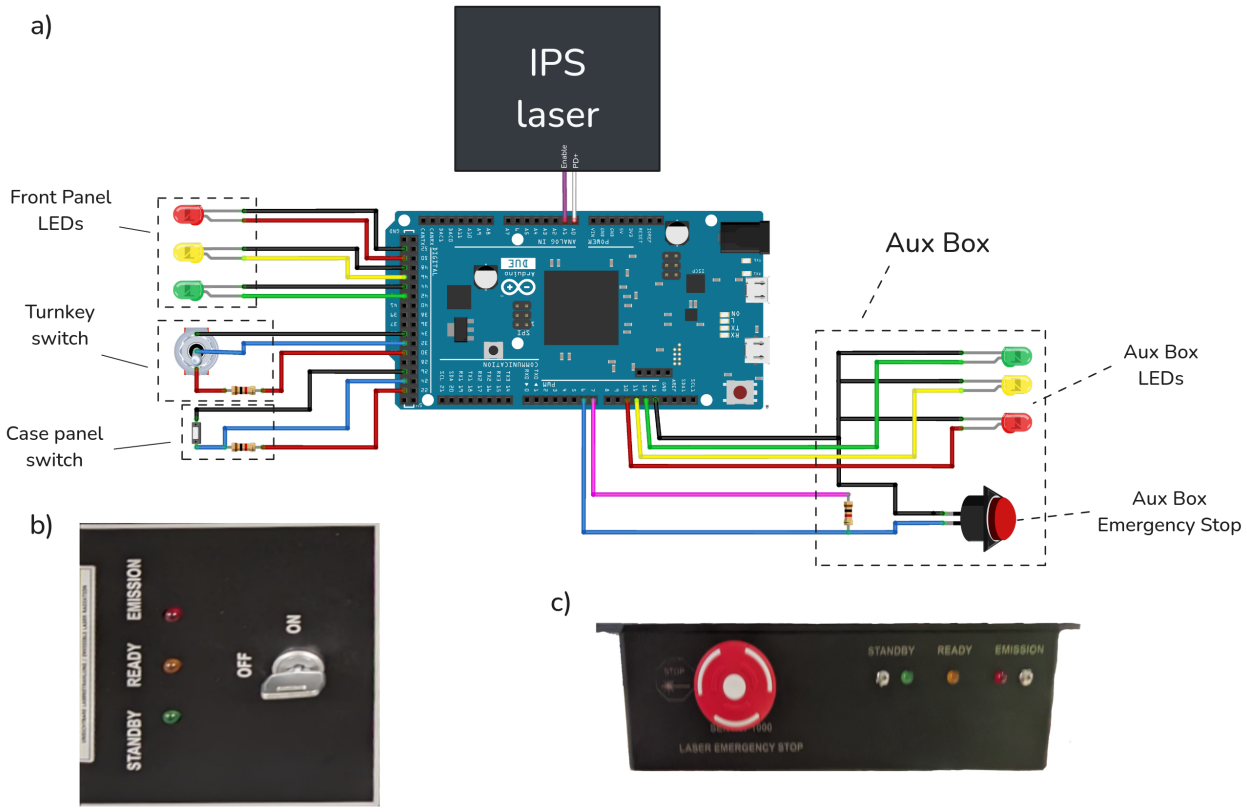


FIGURE 4.11 Schematic diagram of the Laser Safety Control Board used in the Lumed Raman System. (a) Shows the connections between the Arduino Due and the electronic components. (b) is a picture of the unit front panel, showing the laser state LEDs and the turnkey. (c) is a picture of the auxiliary box, showing the emergency stop button and laser state LEDs.

The control board is an Arduino Due, which was chosen because it is easy to program, can

be powered by the same voltage as the laser (12V) and has numerous IO pins. Although the current design of the system only uses a small portion of the available IO pins of the Due, more will be needed in the future to implement additional safety features. An example of such a function not yet added to the safety board is a check that ensures that a probe is connected to the laser output port before enabling the laser. Furthermore, the Due is one of the rare Arduino boards that possesses 2 digital-to-analog chips. This provides the Due with the capability of outputting an analog continuous voltage (0 to 3.3V) without the need for an additional third-party shield. Such a voltage could be used to control older models of IPS lasers or a laser from another manufacturer in analog mode with minimal hardware modifications of the system.

The Due board is connected to three controls and two series of LED indicators. The controls are a turnkey switch and a push button on the main unit front panel, and a large emergency shutoff button on the auxbox. The indicators LEDs are arranged in two groups of three : one red, one yellow and one green that indicate that the laser is emitting light (*emission*), armed (*ready*) and powered (*standby*) respectively. Once powered, the board continuously runs a sequence of checks and adjusts the indicator LEDs and the laser state accordingly. This sequence goes as,

1. The system state is initiated as *standby* (`systemState = 0`)
2. Safety check 1 : that the main unit case is closed
3. Safety check 2 : that the turnkey is in the ON position
4. Safety check 3 : that the emergency shutoff button is not pressed
5. If all safety checks (1 to 3) are satisfied, the laser is enabled (armed/ready, `systemState = 1`) otherwise, it is disabled (standby, `systemState=0`)
6. Emission check : does the voltage on the `PD+` pin of the laser indicates emission of light
7. If emission is detected, `systemState=2` (emission)
8. The LED voltages are adjusted to reflect the system state
   - `systemState==0` : standby LEDs ON, others OFF
   - `systemState==1` : ready LEDs ON, others OFF
   - `systemState==2` : emission LEDs ON, others OFF

The safety check sequence is implemented in C and compiled and uploaded to the Due board via the Arduino IDE software. Each loop of the sequence takes 50 ms when reporting the system state to a computer via serial communication at a Baud Rate of 9600 and is much faster otherwise. Table C.1 gives a summary of the functions implemented in the safety board

code, their inputs and outputs, and their descriptions. The complete source code (Arduino sketch) of the safety board is available as part of the ORAS source code.

4.4 The Open Raman Acquisition Software

The second major part of the Lumed Raman Platform is the software responsible for controlling the different instruments part of the Lumed Raman System (i.e. laser, camera and powermeter) and providing the operator with an easy-to-use Graphical User Interface (GUI) for taking measurements and displaying results.The Open Raman Acquisition Software (ORAS) was created for that purpose. In contrast with the previous software programs developed and used by the Lumed Lab, ORAS needed to be modular, easier to upgrade and maintain and implement machine learning functionalities. As such, the software architecture was redesigned using TimByte [146] as a starting point, but implemented in Python instead of Matlab. The platform's software architecture is illustrated in Figure 4.12. It is broken down in 4 separate python packages, `ORAS`, `lumed_andor`, `lumed_ips` and `lumed_tpm`. `ORAS` is the acquisition software and interfaces with the other packages to manage the instruments, take acquisition, handle spectrum processing, use machine learning models and export data. A list of the different modules that comprise ORAS and of their descriptions is presented in Table 4.7.

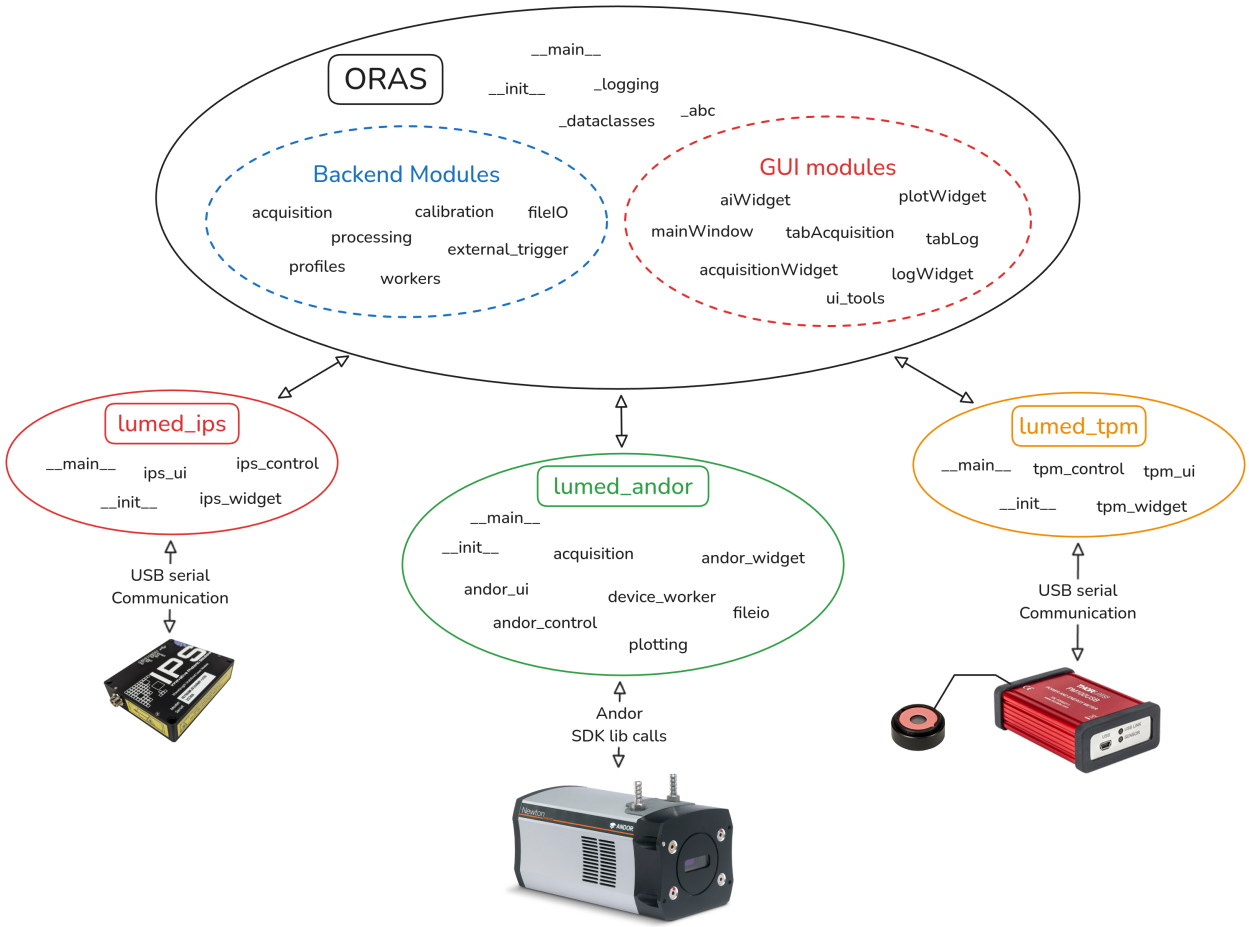


FIGURE 4.12 Overview of the Lumed Raman Platform's software architecture, illustrating the modular structure of the system, which includes the ORAS acquisition software and how it interfaces with the supporting packages : `lumed_andor`, `lumed_ips`, and `lumed_tpm`.

TABLE 4.7 List of modules comprising the Open Raman Acquisition Software (ORAS) with their descriptions, outlining the functionality and purpose of each module within the system.

ORAS modules	description
oras.__main__	the entry point of the software. It initializes the GUI and loads the dependencies.
oras.__init__	initialization of the code and package level imports.
oras._abc	defines a collection of Abstract Base Class for interfacing with data models and external dependencies.
oras._dataclasses	defines the dataclasses used throughout the ORAS codebase.
oras._logging	definition and setup of the logger.
oras.tabAcquisition	definition and setup of the Graphical User Interface GUI for the acquisition tab.
oras.tabLog	setup of the logging tab.
oras.backend.acquisition	definition of the `RamanAcquisition` class and acquisition sequence.
oras.backend.calibration	features related to calibration of the laser power curve.
oras.backend.external_trigger	module that handles triggering and communication between ORAS and other applications over network sockets.
oras.backend.fileIO	module that implements file import and export.
oras.backend.processing	module that implements spectrum processing via `ORPL`.
oras.backend.profiles	module that implements acquisition profiles import and export.
oras.backend.workers	module that defines the base pyqt5 worker class used for threading and parallel execution.
oras.ui.acquisitionWidget	definition of the ui class for the acquisition widget.
oras.ui.aiWidget	definition of the ui class for the ai .
oras.ui.logWidget	definition of the ui class for the logWidget (main widget of the logging tab).
oras.ui.mainWindow	definition of the ui class for the main window of ORAS.
oras.ui.plotWidget	definition of the ui class for the plot widget.
oras.ui.ui_tools	implementation of various miscellaneous tools to help with ui setup.

The 4 packages part of the LRS software platform are hosted on the Lumed Lab GitHub page (https ://github.com/lumed-mtl) and distributed under GPL. They have been developed and tested for use on Linux with a Python v3.12 environment. The recommended operating system for ORAS is Linux Mint v22 or newer, for which an installation guide is presented in Appendix B. The framework used to create the Graphical User Interface of ORAS and the instrument control widgets is QT and is accessed through the PyQT5 bindings. A list of the dependencies used throughout the LRS software platform and their respective licenses is presented in Table 4.8.

TABLE 4.8 – List of software dependencies used across the LRS software platform, including associated licenses.

Dependency	Description	License	PyPi Link
`dacite`	creation of data classes from dictionary	MIT	dacite
`dill`	data serialization	BSD-3	dill
`joblib`	provides lightweight data pipelining in python	BSD-3	joblib
`matplotlib`	plotting and visualization of data in python	PSFL	matplotlib
`numpy`	scientific computing with Python	BSD, numpy	numpy
`orplib`	Open Raman Processing Library	MIT	orplib
`psutil`	cross-platform library for retrieving information on running processes and system utilization	BSD-3	psutil
`pyperclip`	cross-platform Python module for copy and paste clipboard functions	BSD	pyperclip
`PyQt5`	python bindings for QT5	GPL-3	PyQt5
`pyqt5-fugueicons`	provides the icons in the collection Fugue Icons created by Yusuke Kamiyamane	MIT	pyqt5-fugueicons
`pyserial`	Python Serial Port Extension	BSD	pyserial
`pyusb`	USB devices communication in Python	BSD	pyusb
`PyVISA-py`	backend that implements a large part of the "Virtual Instrument Software Architecture" (VISA) in pure Python	MIT	PyVISA-py
`qtmodern`	provides QT with additional themes (light and dark modern)	MIT	qtmodern
`scikit-learn`	module for machine learning built on top of SciPy	BSD	scikit-learn
`tomli`	library for parsing TOML	MIT	tomli
`tomli-w`	library for writing TOML	MIT	tomli-w

Continued on next page

Continued from previous page			
Dependency	**Description**	**License**	**PyPi Link**
`zeroconf`	A pure python implementation of multicast DNS service discovery	LGPL	zeroconf

In addition to ORAS, the LRS software platform relies on stand-alone packages that handle device communication and control. This architecture was chosen instead of implementing device control within the ORAS codebase because it limits coupling between instruments and improves cohesion of the code. As it is, all instrument control packages follow the same basic structure and are comprised of at least

— a module that implements a single class for the control of the instrument (`XYZ_control`),
— a module that set up the GUI elements and layout into a custom QT widget (`XYZ_ui`) and
— a module that implements the GUI internal logic and connects the elements together (`XYZ_widget`).

Following this structure, the instrument control class is responsible for the abstraction of the communication and control mechanisms specific to each instrument. Then, the widget can be embedded into a bigger application (i.e. ORAS) to which an instance of the instrument control class is passed for back-end control while the embedded instrument widget exposes manual control of each device. Details pertaining to the modalities of each of the instruments are presented next.

4.4.1 The Andor Camera Module

Communication and control of Andor cameras is achieved via the Andor Software Development Kit (SDK). At the core of the Linux version of the SDK is a shared library (i.e. `.so` file equivalent to a Windows `.dll`) that exposes camera functionality to a range of programming environments, including, C, C++, C#, Visual Basic and LabVIEW. The shared library is loaded in a python environment by using the Python `ctypes` built-in module and is wrapped inside the `AndorCamera` class implemented in the `andor_control` module. This pure python class acts as a layer of abstraction between the Andor widget or other python applications and the shared library that controls the camera. Figure 4.13.a shows an example where an instance of the `AndorCamera` class is used in a python shell, and Figure 4.13.b shows the Andor camera widget running in a standalone QT application.

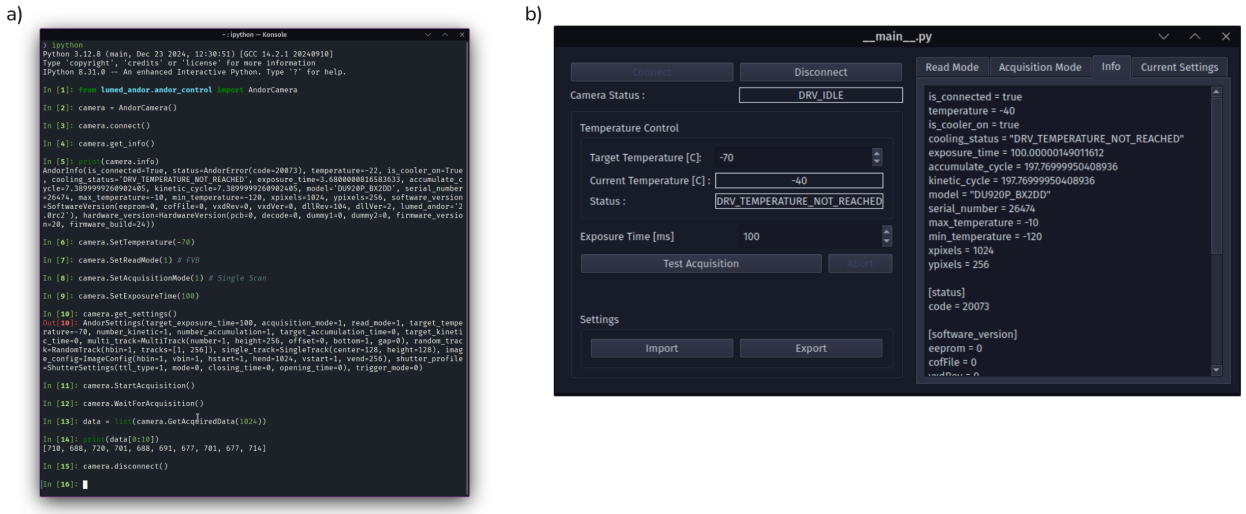


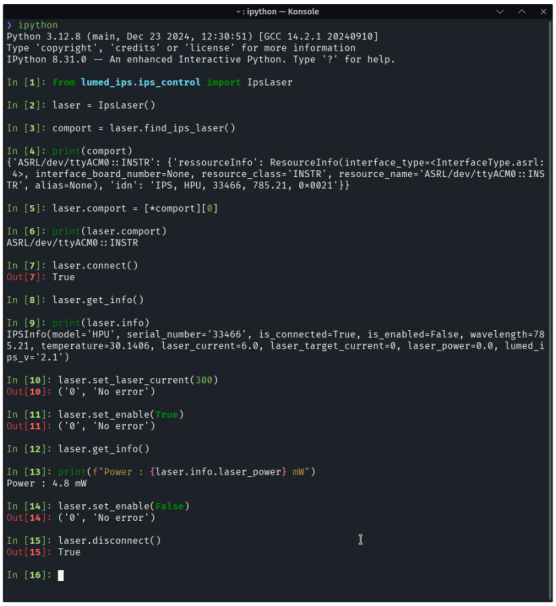
FIGURE 4.13 (a) Example of the `AndorCamera` class used in a Python shell to control the camera. (b) Screenshot of the Andor camera widget in a standalone Qt application.

In addition to `andor_control`, `andor_ui` and `andor_widget`, `lumed_andor` also comprises additional modules that do not have equivalents in the `lumed_ips` and `lumed_tpm` packages. They are the `acquisition`, `device_worker`, `fileio` and `plotting` modules. These were implemented because operation of the Andor camera requires more in-depth and complex steps when compared to the IPS laser or Thorlabs Powermeter. As such, these additional modules provide utilities to simplify control and set-up of acquisitions, low-level threading, reading and writing device configuration files and data visualization of camera acquisitions.

4.4.2 The IPS Laser Module

The digital U-type module lasers from IPS are controlled via serial commands that follow the Standard Commands for Programmable Instruments (SCPI). This means that no additional drivers or third-party libraries are required to control the laser. Instead, serial communication at a baud rate of 115200 is established, and plain text commands are sent to the device via USB. Similarly, as for the Andor camera, the `ips_control` module implements `IpsLaser`, an instrument control class that serves as an abstraction layer between the IPS widget or other python applications and the laser. Figure 4.14.a shows an example where the laser is controlled via a python shell, and Figure 4.14.b shows the IPS widget running in a standalone QT application.

a)



b)

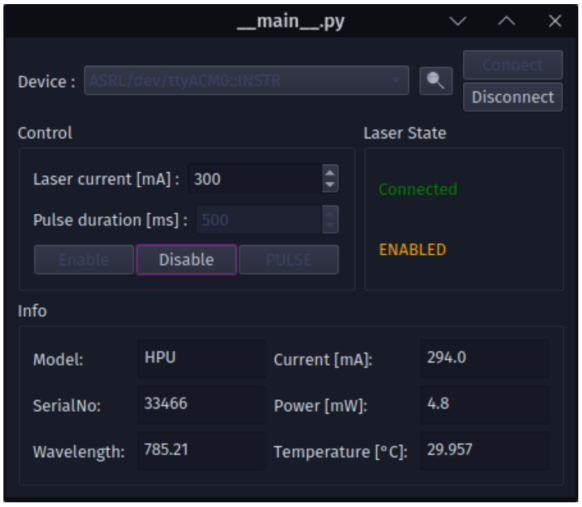
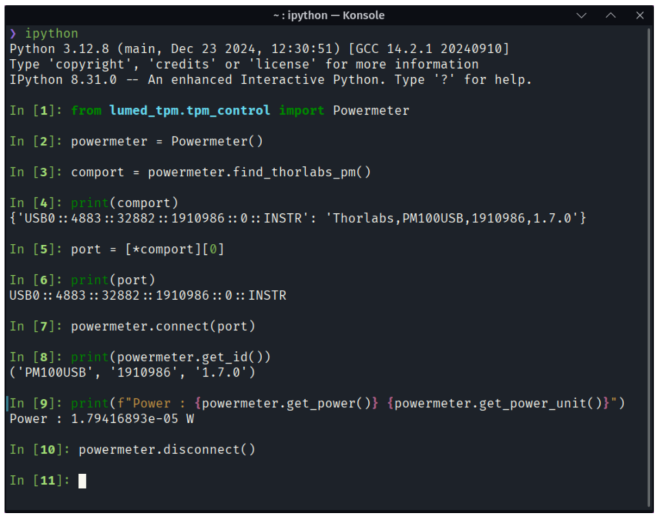


FIGURE 4.14 (a) Example of the `IpsLaser` class from the `ips_control` module within a Python shell. (b) Screenshot of the IPS laser control widget in a standalone Qt application.

4.4.3 The Thorlabs Powermeter Module

As for the IPS laser, the Thorlabs powermeter is controlled with SCPI serial commands sent in plain text via USB. The `Powermeter` class implemented in the `tpm_control` module from the `lumed_tpm` package is used to control the powermeter to take power measurements at a specific wavelength. Figure 4.15.a) shows an example of this in a python shell, and Figure 4.15.b shows the powermeter widget running in a standalone QT application.

a)



b)

FIGURE 4.15 (a) Example of the `Powermeter` class from the `tpm_control` module within a Python shell. (b) Screenshot of the Thorlabs powermeter control widget in a standalone Qt application.

4.5 ORAS's Graphical User Interface

When first launching ORAS, the user is presented with the main window of the software and the three main tabs (Figure 4.16). The *Devices* tab contains the *Summary* panel, where the user can connect and disconnect the various instruments and monitor their respective status. The *Raman Acquisition* tab is where the operator configures and takes acquisitions and where the spectra and classification outputs are displayed. The *Log* tab contains a detailed timestamped log of the current software session, which can be copied to clipboard or accessed in the ORAS log directory.

FIGURE 4.16 Main GUI of the ORAS software, showing the three primary tabs : Devices for instrument connection and status monitoring, Raman Acquisition for configuring and taking spectral measurements with live display and ML predictions, and Log for accessing a timestamped session history.

The user can change the software mode by interacting with the dropdown selection menu located in the top-right corner of the main window. The different modes are the *Operator mode*, the *Scientific mode* and the *Developer mode* which provide increasing level of access to the different features and coincidentally fewer safeguards throughout the software. The *Operator mode* is selected by default and is intended for measurements in patient studies. It provides the user with every feature necessary to take Raman measurements with minimal configuration of the instruments, where safety and software stability are paramount. Putting the software in *Scientific mode* unlocks additional panels along the *Summary* panel

in the *Devices* tab (Figure 4.17). These newly accessible panels contain embedded versions of the instrument control widgets described previously and provide full manual control to the user. This can be used to explore and test non-conventional acquisition settings for the instruments, such as putting the camera in a different read or acquisition mode or manually enabling/disabling the laser. *Developer mode* bypasses all safeguards in the software and is only advisable for troubleshooting and software development, as it can easily lead to software crashes.

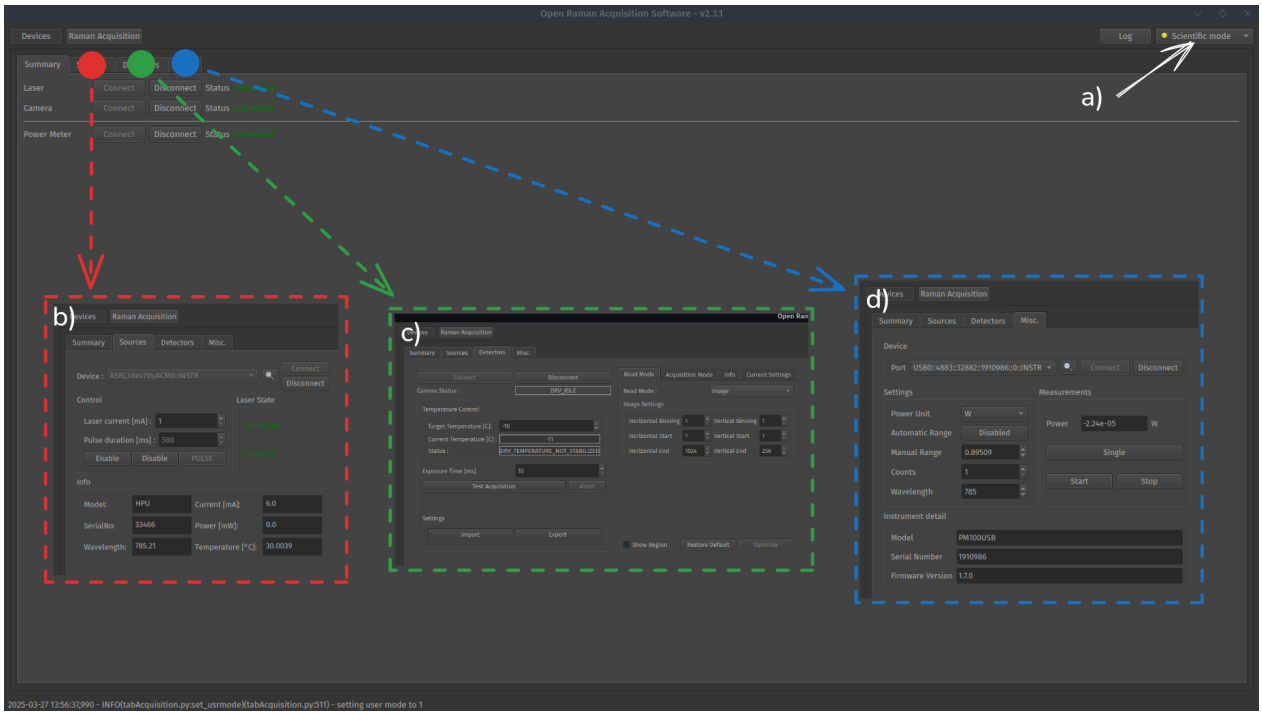


FIGURE 4.17 ORAS's *Devices* tab showing instrument control widgets when the software is put in *Scientific mode*. (a) is the mode selection dropdown menu. (b) is the laser control widget. (c) is the camera control widget. (d) is the laser control widget.

The *Raman Acquisition* tab (Figure 4.18) is where the user configures acquisitions settings, takes measurements and views results. The left sidebar contains the *System Status* panel that displays the instrument status, the currently loaded acquisition profile, ML model and calibration curves. Below is the *Data Export Settings* panel, where the user configures the export location and filename for acquisitions. The *Acquisition Profile*, *Machine Learning* and *Calibration* panels provide the user with means to export and import various profiles and calibration data. Finally, the *Acquisition Control* panel contains a dropdown menu that is used to select the currently loaded acquisition, an interactive text box to view and modify comments and annotations and a start and abort button to launch and cancel acquisitions.

The right sidebar appears when putting the software in *Scientific mode*. It contains various controls that are used to configure the acquisition. The user can export the current configuration as a new acquisition profile for later use. The middle section contains the ML output widget and a plot widget that displays acquisition results. The information displayed in the ML output widget is dynamic and is adjusted upon loading a new model to match its outputs (ie. various prediction classes, scores, probability, ...). In the example shown in Figure 4.18, the model loaded is a toy model 4-way classifier that displays the Raman spectrum quality factor (ASSI) and the individual posterior probability of the 4 one-vs-all sub-classifiers, along with a confidence index for the final result.

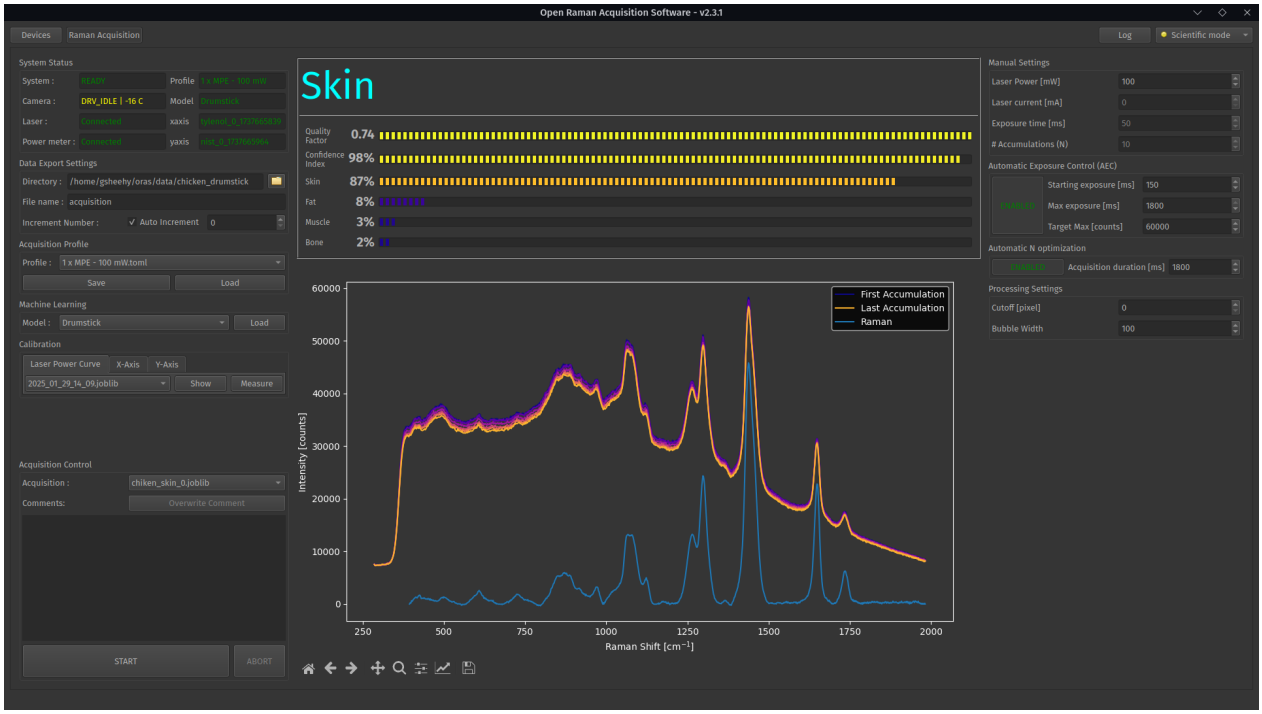


FIGURE 4.18 Overview of the Raman Acquisition tab in ORAS. The interface includes panels for system status, data export, acquisition profile management, calibration, and manual control. Measurement results are displayed in real time via a spectral plot and a dynamic machine learning output widget, which adjusts based on the loaded model.

4.5.1 Acquisition Sequence

The acquisition sequence of ORAS is launched every time the user takes a measurement by pressing the *Start* button. The sequence begins by configuring the laser and the camera with the settings specified in the loaded acquisition profile. This includes laser power, exposure time and number of repeated measurements. After the acquisition sequence is completed, the series of repeated measurements (raw accumulations) from the camera is returned to the

computer. The measured spectra are then processed using ORPL as per the workflow outlined in the previous chapter and displayed in the plotting widget. However, if a Machine Learning model is loaded, the raw accumulations are instead passed to the processing sequence of the loaded model. The model outputs and processed spectra are displayed in their respective widgets afterward. The user can configure an acquisition by manually specifying the exposure time and the number of repeated measurement, or by relying on the Automatic Exposure Control (AEC) and the Automatic N Optimization (ANO). With the ANO enabled, the user specifies the total duration for the acquisition instead of the number of repeated measurements. The system will automatically take as many repeated measurements within the time window specified without going over. With the AEC enabled, the user no longer needs to set the exposure time for a measurement and instead configures the target for the intensity level (counts) to be reached. The AEC will optimize the exposure time necessary to reach that target for each acquisition. The details of this procedure are presented in the following paragraphs.

A common acquisition sequence is represented in Figure 4.19 as a timeline of events. The example is shown with the AEC and the ANO enabled. This is the intended mode of acquisition for ORAS in most scenarios. The sequence begins with an instrument configuration step. The laser current (in mA) is set so the emission output power (in mW) at the tip of the instrument matches the value specified in the acquisition profile. The current-power curve of the laser measured during system calibration is used for the conversion between laser current (mA) and output power (mW). The camera's acquisition mode is set to "Fast Kinetic" and the exposure time and number of repeated measurements are determined with the AEC and ANO routines. A series of three short-duration spectra is measured and returned to the computer. The exposure time is configured to fill the dynamic range of the camera CCD array, and the number of repeated measurement is set to fit as many distinct exposures within the total acquisition duration specified in the ANO settings. The acquisition sequence continues with the measurement of a single background spectrum, during which the laser remains disabled. This background spectrum is returned to the computer and saved for use in processing and analysis. The main acquisition finally takes place. The laser is enabled, and the N repeated measurements are taken by the camera. The laser is disabled and the data is returned to the computer one last time as the sequence completes.

The sequence depicted in this example is representative of an acquisition taken using the `2 x MPE - 150 mW` acquisition profile included in ORAS. With this profile, the AEC and ANO are enabled, the maximum exposure duration is of 2700 ms and the starting exposure for the three AEC spectra is of 150 ms. The only duration that could significantly differ between the example presented and a real-world measurement using this profile is the exposure time for

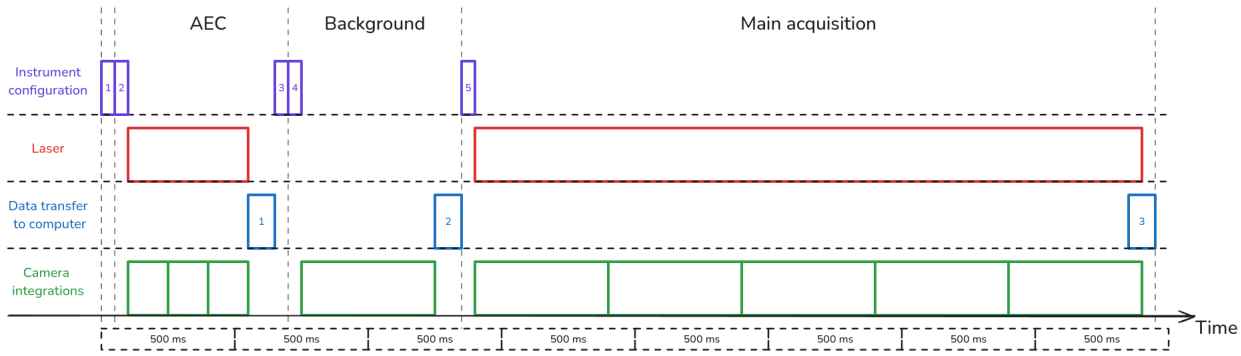


FIGURE 4.19 Timeline of a typical ORAS acquisition sequence with Automatic Exposure Control (AEC) and Automatic N Optimization (ANO) enabled. The process includes instrument configuration, preliminary short acquisitions for exposure calibration, background measurement without laser excitation, and the final acquisition sequence with N repeated measurements determined within the specified acquisition duration.

the background measurement. In this scenario, the duration for this step could theoretically be increased up to 2700 ms which corresponds to the maximum exposure time for a single accumulation. The duration of the instrument configuration and data transfer depicted in the timeline are conservative overestimates of the real-world duration for these steps. Instrument configuration takes on the order of a few milliseconds, and data transfer only approximately reaches 100 ms when taking full sensor images with the cameras. Considering this, the total duration of the acquisition presented in this example would approximately be of 4 seconds.

4.5.2 Automatic Exposure Control

The Automatic Exposure Control (AEC) is an optional sub-routine that runs at the beginning of the acquisition sequence. Its purpose is to optimize the exposure time so that the camera's dynamic range is filled for each spectrum of the series of repeated accumulation that will follow. This is achieved by taking a series of three quick spectra (with laser enabled), identifying the maximum intensity, and scaling the exposure time with the following formula,

$$t^\star = \min\left(\frac{I^\star}{I_{AEC}} \times t_{AEC},\ t_{max}\right) \tag{4.10}$$

$t^\star$ is the optimized exposure time that will be set for the following accumulations, $I^\star$ is the target intensity, t_{max} is the *maximum exposure* time, I_{AEC} and t_{AEC} are the maximum intensity and the exposure time (*starting exposure*) for the three AEC spectra. Three spectra are measured to make the AEC routine robust even in the presence of cosmic rays by taking the median across the three intensity values for each wavelength. All parameters involved

with the AEC are summarized in Table 4.9 and a flowchart of the algorithm is illustrated in Figure 4.20.

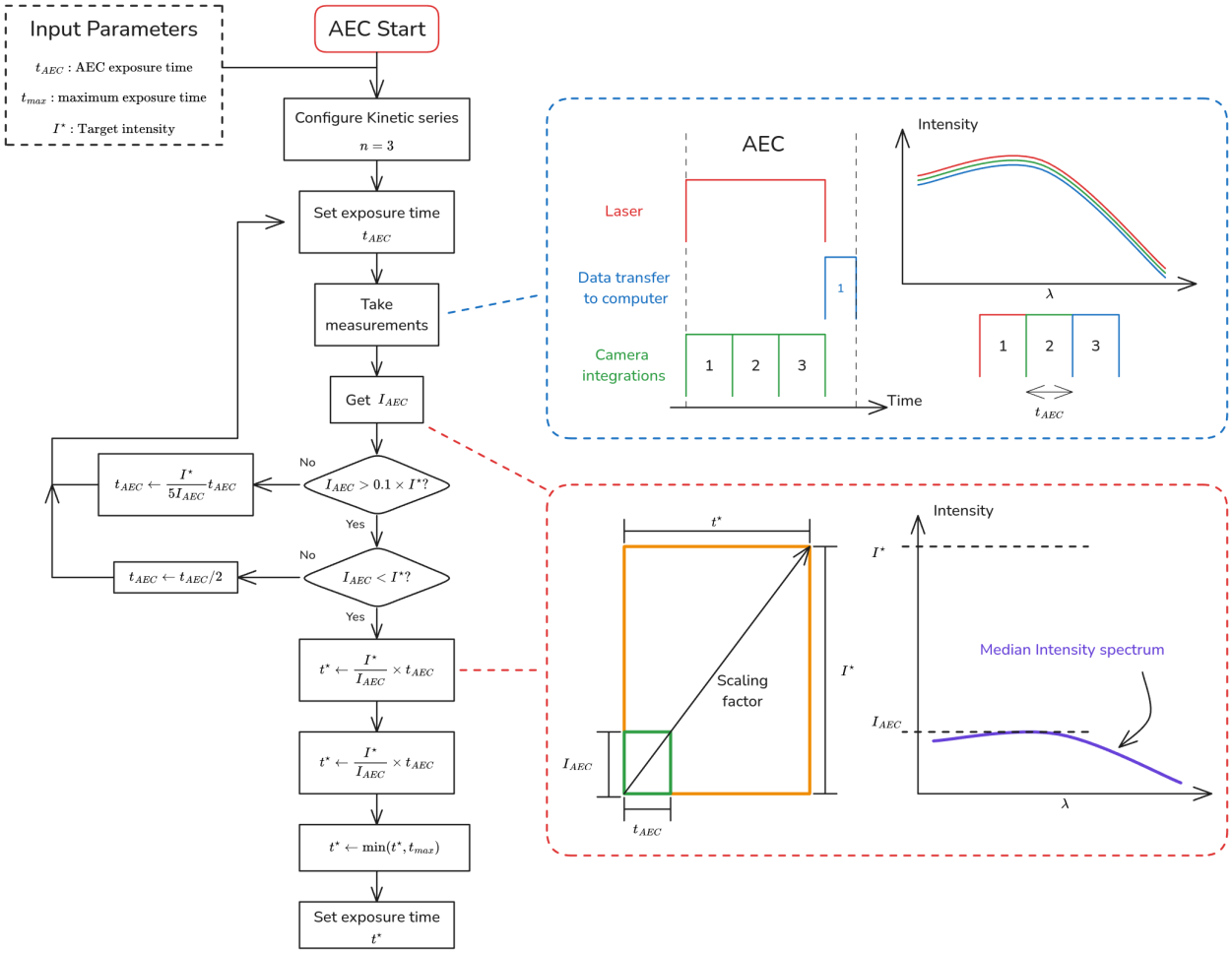


FIGURE 4.20 Flowchart of the Automatic Exposure Control (AEC) algorithm used in ORAS. The procedure estimates an optimal exposure time by acquiring three preliminary spectra and applying a scaling formula to set the exposure time for the subsequent accumulations.

TABLE 4.9 Summary of parameters used in the Automatic Exposure Control (AEC) sub-routine.

Parameter	**Symbol**	**Origin**	**Description**
AEC exposure time (starting exposure)	t_{AEC}	configured	The exposure time for the three AEC spectra
AEC max intensity	I_{AEC}	measured	The maximum intensity of the three AEC spectra
Optimized exposure time	$t^{\star}$	Equation 4.10	The optimal exposure time determined by the AEC sequence that will result in the target intensity
Maximum exposure time	t_{max}	configured	The maximum limit for the optimized exposure time
Target intensity	$I^{\star}$	configured	The desired target for the maximum intensity of the measurements that follow the AEC

The configurable parameters for the AEC are the *starting exposure*, the *maximum exposure* and the *target max* and are visible in the right sidebar of ORAS (Figure 4.18). The target max ($I^{\star}$) is chosen to be as close as possible to the saturation intensity of the system's camera to fill the dynamic range (16-bit digitization, $2^{16} - 1 = 65,535$ counts). However, most cameras will start clipping the signal at a slightly lower intensity count. As such, it is best to account for a small buffer and aim for a target intensity of $60,000$ counts and adjust for each system's camera. The *max exposure* parameter serves as the hard-limit for the optimized exposure and cannot be exceeded by the AEC optimization sequence. If the optimized exposure calculated during the AEC sub-routine exceeds this limit, the exposure time for the following acquisition will be set as the maximum exposure time and a single accumulation will be measured. The configured AEC exposure time (starting exposure) should be long enough for the spectrum intensity to exceed the noise floor of the camera, but short enough to prevent saturation. If it is too short and the measured intensity is smaller than 10 % of the target intensity ($I_{AEC} < 0.1 \times I^{\star}$), the AEC exposure time is increased. If it is too long and the measured intensity is greater than the target ($I_{AEC} > I^{\star}$), the AEC exposure time is halved. In either case, the AEC sequence is repeated with the adjusted exposure until it reaches a suitable value for the sample being measured. Because of this, adjusting the AEC starting exposure time to a rough estimate of the same order of magnitude as the optimized exposure time avoids unnecessary steps in the AEC sequence. The AEC starting exposure set by default in ORAS is 150 ms and is the recommended value for the Lumed systems when measuring biological specimen in clinical studies.

4.5.3 Acquisition Profiles

Acquisition settings and instrument configurations are stored within ORAS as instances of the `AcquisitionProfile` python dataclass defined in the `_dataclass` module. User interactions with the controls in the Raman Acquisition tab change the values stored in the instance of the acquisition profile currently loaded. Profiles can be exported from the current configuration or loaded from a `.toml` file that matches the required structure outlined in the ORAS documentation. By default, ORAS includes several profiles intended for use with biological samples, such as clinical studies with patients. These profiles are configured for different levels of exposure as multipliers of the skin MPE. Table 4.10 presents the list of ORAS included profiles and Figure 4.21 illustrates how the profiles compare to the 1×, 2× and 5× MPE limit curves. Each profile makes use of the AEC and ANO to optimize exposure time, but uses different laser power and acquisition durations. The choice of which profile to use depends on the context and environment where acquisitions are made. The most suited profiles for use in clinical studies that involve a surgeon manipulating the probe or instrument are the `1 x MPE - 100 mW.toml` or the `2 x MPE - 150 mW.toml` profiles. While both provide quick measurements that take less than a few seconds, the former involves fewer risks of tissue damage and the latter yields higher SNR spectra. The recommended profile for medical studies where the higher exposure levels are approved by medical personnel with the Lumed systems is the `2 x MPE - 150 mW.toml`. The $5 \times$ MPE profile is not intended for use cases that involve patients or *in-vivo* measurements, as levels of emissions can be unsafe without protection and appropriate safety measures. It is best suited for use where the instrument can be fixed in place within an enclosure, such as with *ex-vivo* studies or with biofluids. In practice, the choice of acquisition settings constitutes a tradeoff between signal quality and risks of tissue damage.

TABLE 4.10 List of dependencies used in the lrs software platform along with their descriptions and licenses.

Acquisition Profile	MPE factor	Acquisition duration	Laser power
`1 x MPE - 50 mW.toml`	1×	4500 ms	50 mW
`1 x MPE - 100 mW.toml`	1×	1800 ms	100 mW
`2 x MPE - 100 mW.toml`	2×	4500 ms	100 mW
`2 x MPE - 150 mW.toml`	2×	2700 ms	150 mW
`5 x MPE - 150 mW.toml`	5×	9000 ms	150 mW

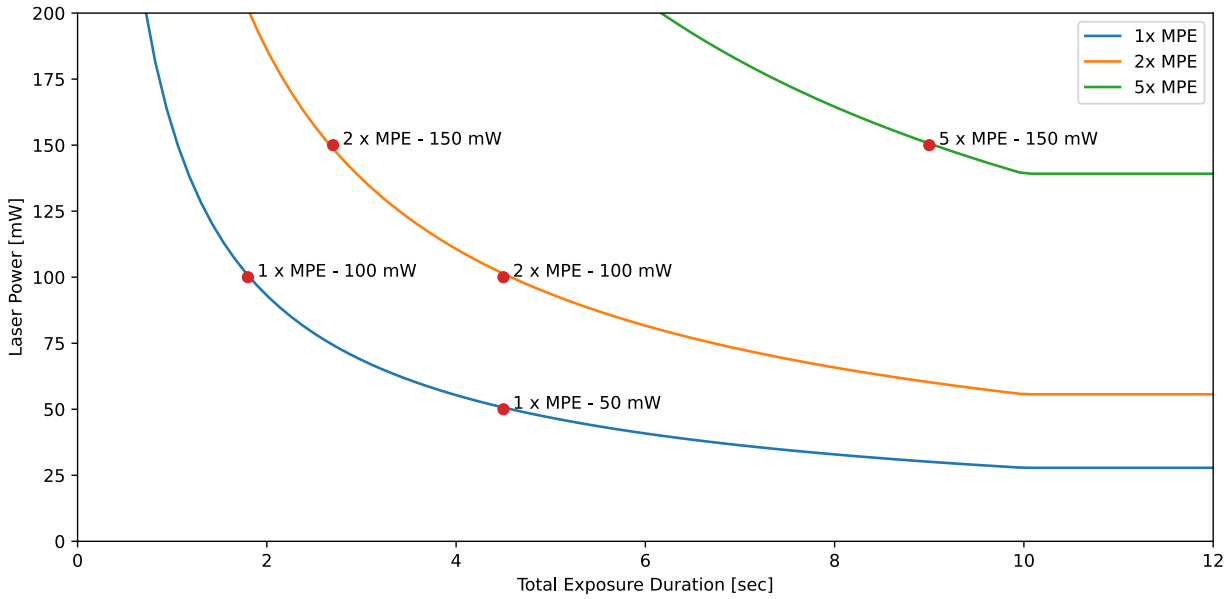


FIGURE 4.21 Comparison of default ORAS acquisition profiles with skin Maximum Permissible Exposure (MPE) curves at 1×, 2×, and 5× multiplication factors.

4.6 Results

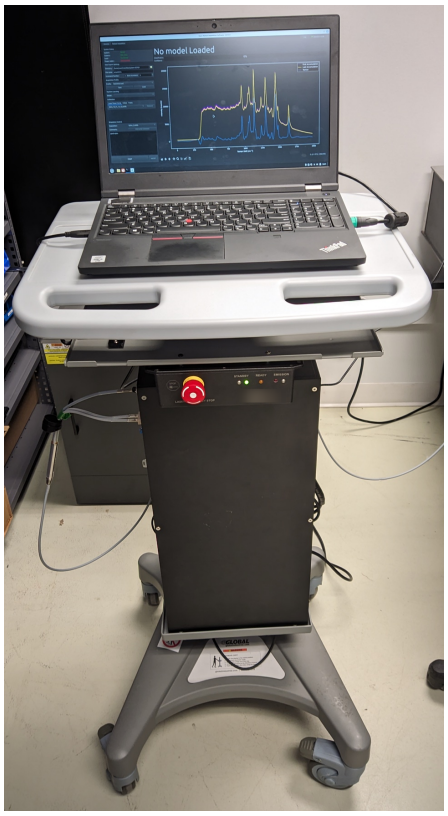
4.6.1 The Lumed System Fleet

Following the description of the system development from the previous sections, and the system assembly guide outlined in Appendix A, multiple Raman spectroscopy systems have

been assembled to support both clinical research and continued platform development. Four complete systems were built since 2024 with the specific objective of acquiring Raman signals from biological tissues *in-vivo*, under clinical study protocols (Figure 4.22b). These systems were designed to comply with hospital safety requirements and to integrate into clinical workflows with minimal disruption. In addition to these, a smaller, more portable unit — referred to as LRS Mini — was developed to facilitate laboratory-based testing and ongoing hardware and software refinement and development (Figure 4.22d). Table 4.11 presents an overview of the clinical systems, collectively referred to as the Lumed System Fleet, including their current deployment location across hospital centres in Montreal and their respective roles in active clinical studies as of April 2025.

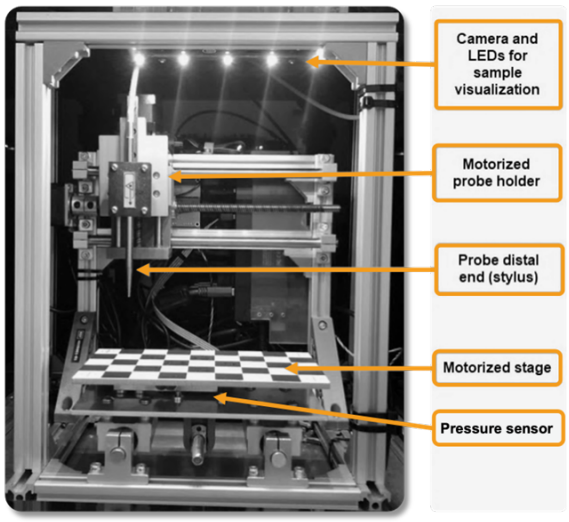
TABLE 4.11 Overview of the Lumed Raman System Fleet, including system designation, add-ons, current deployment location, and associated clinical role, or primary biological sample types targeted in their respective studies.

System designation	**Add-ons**	**Deployment Site**	**Clinical study / Sample types**
Lumed 1	mobile cart, auxiliary box	MUHC network (MNI and Montreal Children's Hospital)	In-vivo Raman spectroscopy during adult brain surgery for the characterization of healthy cortex, gliomas, meningiomas, bulk tumour tissue, and infiltrative cancer.
Lumed 2	mobile cart, auxiliary box	CHU Sainte-Justine	In-vivo Raman spectroscopy during pediatric brain surgery for cancer and Focal Cortical Dysplasia (FCD). Tissue type includes cortical *gyri* and *sulci*, *corpus callosum*, and *fornix*.
Lumed 3	scanner	CRCHUM	large mammal studies (e.g. lamb and pork), transsphenoidal pituitary adenoma surgery [5], transperineal prostate cancer surgery and various other ex-vivo and in vitro samples (e.g. ovarian cancer cell lines and connective tissue sarcoma cell lines)
Lumed 4	mobile cart, auxiliary box, scanner	MUHC Glen site	Lumpectomy specimen excised during breast conserving surgery. Samples may include healthy breast tissue, adipose tissue, lymph ducts and various malignant breast tissue. [48]



(a) The LRS mounted on a mobile clinical cart

(b) The LRS in use during an intraoperative brain measurement in the operating room.



(c) Optional 3D scanner enclosure used for automated probe positioning. Key components are annotated.

(d) The LRS-mini system, a compact and portable variant designed for laboratory-based testing and development.

FIGURE 4.22 Photographs of the Lumed Raman systems and associated hardware.

Of the four clinical systems constructed, three are mounted on wheeled mobile carts and include the auxiliary box that houses the emergency stop button and laser emission state LEDs (Figure 4.22a). Two of the systems — one equipped with the auxiliary box and one without — are further outfitted with a custom-built 3D scanning add-on enclosure (Figure 4.22c). This scanning add-on is adapted from commercial 3D printer hardware, wherein the original printhead has been removed and replaced by a Raman probe mount fixture. The system enables automated positioning of the probe and the acquisition of Raman measurements across a defined sampling grid on specimens placed within the scanner's enclosure. Designed as a plug-and-play module, the scanner can be seamlessly integrated with the Lumed Raman System without requiring hardware or software modifications to the core platform. The probe mounted on the scanner carriage can be readily disconnected from the main unit, allowing it to be replaced with a handheld probe for conventional manual use. This modularity facilitates operational flexibility between automated and manual acquisition modes. While the scanner control software remains under active development in parallel projects, essential connectivity functions have already been integrated into ORAS through the `external_trigger` module. This module exposes a connection API via network sockets — currently restricted to `localhost` connections — that permits external software to initiate measurements, exercise limited control over acquisition parameters, and receive acquired spectral data in real time.

4.6.2 Comparison with the ODS/Reveal Sentry system

To assess the signal acquisition performance of the Lumed Raman System in comparison with a state-of-the-art reference platform, a retrospective analysis was conducted using Raman spectroscopy data acquired during intraoperative measurements in clinical studies. Specifically, data collected with the ODS/Reveal Sentry system was used as a benchmark. While not yet commercially available at the time of this writing, the Sentry system represents, to the best of the author's knowledge, the platform closest to clinical translation and regulatory approval. As such, it serves as a meaningful and appropriate point of comparison for the open-source Lumed Raman System in the context of this work. Notably, the Lumed and Sentry systems share key hardware components (i.e. excitation laser, spectrometer, detector and probe) thereby minimizing instrumental variability. The principal distinctions between the two platforms reside in their respective acquisition software architectures and data collection protocols. This shared instrumentation, coupled with divergent system control and data handling methodologies, further reinforces the significance and interpretability of a comparative analysis focused on signal quality and system performance in clinical applications.

The first step of this analysis involved the aggregation of raw Raman spectroscopy data col-

lected using the Sentry system in the context of multiple brain tumour studies, as summarized in the multicentre study by Ember et al.[56]. These datasets included in-situ Raman spectral measurements obtained from a range of tumour types and healthy brain tissue in patients undergoing surgery at two major institutions (e.g. Montreal Neurological Institute-Hospital[1] and Mount Sinai Hospital, New York[2]).

The second dataset used for comparison originated from a clinical study employing the Lumed Raman System during transsphenoidal pituitary adenoma surgery, as reported by Blanquez-Yeste et al. [5]. In this investigation, a fibre-optic Raman probe specifically designed for endonasal surgical access was employed to perform in situ measurements of intracranial structures relevant to pituitary tumour resection. The study protocol involved the acquisition of Raman spectra directly from exposed tissue surfaces using a sterilizable, angled-tip probe introduced through the nasal cavity during endoscopic procedures (Figure 4.4.c). The protocol included a standardized calibration sequence using both NIST Raman standards[129] and reference compounds, followed by spectral acquisitions performed under low ambient light conditions. Each measurement site was interrogated using a fixed laser power (40 mW) and variable exposure times (optimized with the AEC routine presented earlier), which were adapted to maximize detector usage without inducing saturation. For each site, a series of 10 repeated spectral accumulations were recorded and subsequently averaged. The probe was held in contact with the target tissue by a trained neurosurgery resident, while anatomical localization was guided by preoperative imaging and direct visual assessment. Measurements were acquired from key structures such as the pituitary gland, *sella turcica*, *optic chiasm*, and surrounding *dura mater*. These data, which include raw unaveraged spectral accumulations, were used in the present analysis to evaluate the signal acquisition characteristics of the Lumed Raman System in a clinical intraoperative setting.

To standardize the comparison, the maximum intensity value from the highest individual accumulation (i.e., before averaging) was extracted for each measurement site across both datasets. This metric was selected as a proxy for evaluating the extent to which the signal filled the dynamic range of the camera sensor during acquisition. Because higher signal intensities correspond to improved SNR, this measure was used as a practical indicator of signal quality and acquisition efficiency under clinical conditions. The resulting intensity maxima were then plotted as a function of acquisition index, representing the chronological order of measurements across the combined clinical datasets. This graphical representation allows for a quantitative comparison of acquisition performance between the two systems over time and across distinct clinical contexts.

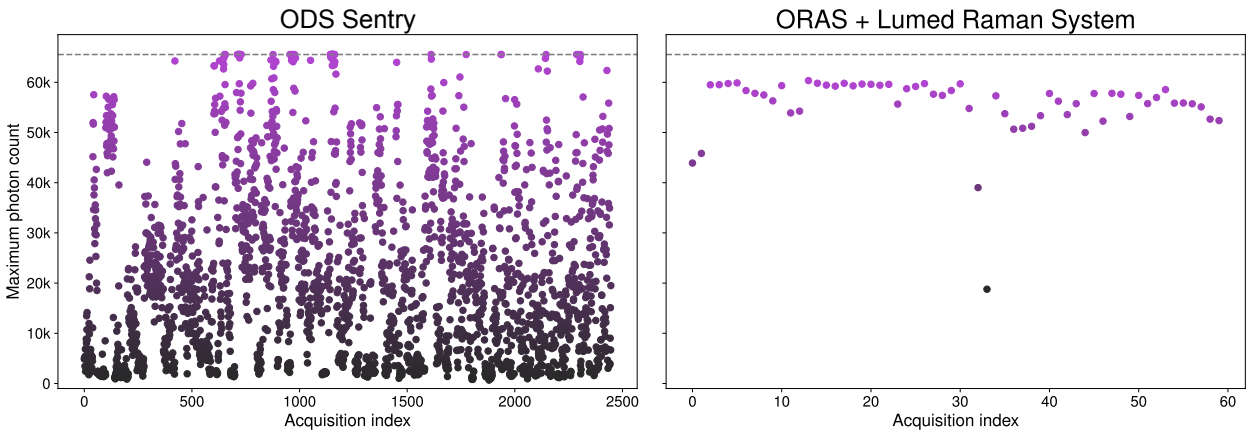


FIGURE 4.23 Maximum intensity values from raw spectral accumulations acquired in vivo using the ODS/Reveal Sentry system (left) and the Lumed Raman System (right). Each point represents the photon count at the signal peak from a single measurement site. The dashed horizontal line denotes the saturation threshold of the camera sensor. Sentry data were collected during brain tumour surgeries at MNI and MSH ; Lumed data were obtained during transsphenoidal pituitary adenoma surgeries.

The distribution of maximum photon counts from the first raw accumulation of each in-vivo measurement is presented in Figure 4.23, comparing the ODS/Reveal Sentry system (left) and the ORAS-controlled Lumed Raman System (right). For the Sentry system, all acquisitions were performed with a fixed exposure time of 100 ms, independent of tissue type or signal strength (2447 acquisitions across 107 patients). This static acquisition parameter prevented dynamic optimization of the sensor's dynamic range. As a result, 77 % of the measurements reached less than 50 % of the sensor's maximum capacity, and 48 % of them filled less than 25 %. At the other end of the distribution, several acquisitions reached full saturation of the sensor, indicating potential signal clipping and loss of spectral integrity.

In contrast, acquisitions performed with the Lumed Raman System were dynamically optimized using the automatic exposure control (AEC) functionality integrated into the Open Raman Acquisition Software (ORAS). A target peak intensity was predefined to avoid saturation while maximizing signal quality. During the clinical study involving 60 acquisition sites from 8 patients[5], the target value was progressively adjusted from 65,000 to 60,000, and ultimately to 58,000 photon counts, to avoid flattening of the spectral curves observed near the upper end of the sensor's range.

The effect of this optimization is clearly visible in the distribution of peak intensities. The

1. Montreal Neurological Institute-Hospital, McGill University, Montreal, QC, Canada.
2. Mount Sinai Hospital, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

data acquired with the Sentry system exhibit substantial variability and inconsistency, with numerous values falling below or exceeding ideal operating conditions. Conversely, the data acquired with the Lumed system are predominantly concentrated within the 40,000 to 60,000 count range, close to the dynamic range target, with no instances of saturation. The only notable deviations are four lower-intensity points : the first two acquisitions of the dataset and two consecutive measurements (indices 32 and 33). The former are attributed to the initial adaptation phase typical when initiating a new clinical study, which often involves calibration of workflow, adaptation to specific surgical conditions, and onboarding of new clinical personnel. Despite these challenges, the signal quality improved rapidly within the first two measurements. This contrasts with the Sentry dataset, where low-count acquisitions frequently occurred across early phases of multiple clinical studies.

The two lower outliers at indices 32 and 33 in the Lumed dataset were explicitly annotated in the ORAS metadata as having occurred in the presence of excessive blood pooling on the tissue surface. Such annotations were entered via the comment field available in ORAS at the time of acquisition. This aids retrospective data interpretation and exemplifies the system's suitability for routine clinical use, where integration of metadata capture and ease of operation are essential for robust intraoperative application.

4.7 Discussion

The Open Lumed Raman Platform represents a comprehensive and modular system for in-vivo Raman spectroscopy tailored to the constraints and requirements of clinical environments. This chapter has detailed the platform's technical architecture, encompassing both hardware and software components, and has outlined the development of a fully operational fleet of four clinical systems currently deployed across multiple hospital sites. Through the integration of flexible acquisition software, hardware safety mechanisms, and optional extensions such as 3D scanning modules, the platform achieves a level of adaptability and reliability necessary for intraoperative use. Comparative results with the ODS/Reveal Sentry system—one of the most advanced pre-commercial platforms—highlight the efficacy of the Lumed system's signal acquisition strategy, particularly its use of dynamic exposure control to optimize signal quality without compromising sensor integrity. Collectively, these elements position the Open Lumed Raman Platform as a robust and open-source alternative for advancing label-free optical diagnostics in clinical research and translational medicine.

A central design objective of the Open Lumed Raman Platform was to ensure maximal modularity and adaptability at both the hardware and software levels. This principle is most clearly reflected in the architecture of the Open Raman Acquisition Software (ORAS), which employs

a minimally coupled structure between its core logic and the control interfaces for individual hardware components. Instrument control is abstracted into discrete Python packages, each implementing a standardized class interface as defined in the ORAS documentation. As a result, the integration of new hardware — such as alternative lasers or camera systems — requires only the implementation of a new device-specific control class, without necessitating significant modification to the central codebase or its user interface logic. Additionally, this architecture promotes software reusability : each device control package is functionally independent and can be deployed as standalone control software outside the ORAS framework. This approach not only streamlines integration across diverse experimental setups but also fosters software interoperability and parallel development across research groups, reducing redundancy and enabling broader collaborative applications. Furthermore, the adoption of the GNU General Public License (GPL) for both the ORAS codebase and the associated instrument control packages reinforces these design principles at the licensing level. By requiring that any derivative works remain open-source, the GPL prevents the future proprietary enclosure of the platform and preserves its long-term accessibility. This ensures that enhancements and modifications made by external developers are also shared under the same terms, thereby fostering a sustainable and collaborative development ecosystem. In this way, the licensing strategy acts as a structural safeguard for the openness, reusability, and extensibility that define the Open Lumed Raman Platform.

The integration of automated subroutines for acquisition control within ORAS — specifically the Automatic Exposure Control (AEC) and Automatic N Optimization (ANO) — has significantly streamlined the process of data acquisition in the clinical setting. These features enable the system to autonomously adjust key acquisition parameters such as exposure time and the number of accumulations to maximize signal quality while preventing sensor saturation or excessive measurement times. As a result, the procedure for collecting high-quality Raman spectra is both rapid and operator-independent, two critical requirements for clinical integration of the platform. From system initialization to the visualization of a Raman spectrum on ORAS's graphical interface — complete with real-time AI-based tissue classification — the process can be completed in fewer than 15 mouse clicks. This degree of automation is essential for ensuring system usability in the time-constrained and high-stakes environment of the operating room. In addition to automation, safety and standardization are central to the acquisition workflow. ORAS supports the use of predefined acquisition profiles that encode both system parameters and safety limits. These profiles are named and structured according to the Maximum Permissible Exposure (MPE) guidelines relevant to laser-tissue interaction, thereby providing clinical users with immediate and intuitive insight into the safety characteristics of each configuration. By abstracting safety-critical parameters into se-

lectable profiles, the system eliminates guesswork during setup and minimizes the risk of user error. This approach ensures that the platform can be safely operated by clinical personnel without specialized training in laser safety or Raman spectroscopy, thereby enhancing both adoption and compliance within the surgical workflow.

The comparison between the Lumed Raman System and the ODS/Reveal Sentry system further underscores the practical advantages of the open-source platform in a clinical context. While the Sentry system demonstrated high classification accuracy in a multicentre study involving over 100 patients—achieving sensitivities and specificities exceeding 90 % across multiple brain tumour types—the fixed-parameter acquisition protocol used in that study limited the signal quality in a substantial proportion of measurements. By contrast, the Lumed Raman System, with its integration of dynamic acquisition optimization routines, produced data with consistently high signal quality, while also eliminating saturation events and minimizing underexposed measurements. The resulting improvements in data quality, consistency, and efficiency directly address several of the key limitations observed in the Sentry dataset. Moreover, the modularity of the Lumed platform and the richness of metadata captured in ORAS further support rigorous labelling and contextualization of each measurement. While the Sentry study serves as an important milestone demonstration of the clinical utility of *in-vivo* Raman spectroscopy, it is reasonable to suggest that similar or even improved classification performance would be achieved using models trained on data acquired under the more controlled and optimized conditions enabled by the Lumed Raman System. This highlights the platform's potential not only as a tool for data acquisition, but also as a foundation for advancing machine learning-driven intraoperative diagnostics.

In summary, the Open Lumed Raman Platform constitutes a versatile, open, and clinically oriented solution for in-vivo Raman spectroscopy. By integrating modular hardware, extensible open-source software, and automated acquisition routines aligned with clinical workflows and safety standards, the platform addresses many of the barriers that have historically hindered the translation of Raman technologies into surgical environments. Its deployment across multiple hospital centres and its demonstrated performance in clinical studies underscore its readiness for broader translational research applications. The comparative advantages over existing systems, particularly in terms of data quality, usability, and interoperability, position the Lumed Raman System as a foundational tool for the next generation of optical diagnostics. As ongoing efforts continue to expand its capabilities and integrate advanced analytical models, the platform is poised to contribute meaningfully to both clinical research and intraoperative decision support.

CHAPTER 5 CONCLUSION AND RECOMMENDATIONS

5.1 Advancement of knowledge

This thesis addresses some of the remaining key challenges associated with the integration of Raman spectroscopy into clinical and translational healthcare environments. Contributions span algorithm development, hardware and software engineering, and clinical system integration, with all components released under open-source licenses to promote transparency, reproducibility, and broader adoption.

At the core of this effort is the intention to create an open, modular framework that supports both the development and application of Raman spectroscopy in biomedical research. Each objective was pursued with a strong emphasis on open-source dissemination, ensuring that the results produced can be reused, adapted, and critically evaluated by others.

Objective 1 aimed to develop and validate an open-source Raman spectral processing library designed to enhance baseline correction and processing performance for biological sample analysis, with a specific focus on *in-vivo* cancer diagnosis. This objective was addressed through the work presented in Chapter 3 and formalized in the release of the Open Raman Processing Library (ORPL) and peer-reviewed publication[147]. The following breakdown details how each of the sub-objectives contributed to the realization of this goal :

— **Open-Source Raman Processing Library** : The ORPL library was developed as a comprehensive and modular Python package providing standardized tools for processing Raman spectra from CCD-based dispersive systems. It includes functionalities for dark signal subtraction, cosmic ray removal, baseline correction, wavenumber calibration, normalization, and spectral smoothing. Released under the permissive MIT license, ORPL ensures full transparency, reusability, and accessibility for the research community. Since its publication, it has been adopted as the standard processing pipeline for all Raman-related projects conducted within the Lumed research group.

— **Novel Baseline Removal Algorithm** : A key contribution of the library is the introduction of a new baseline correction method, *BubbleFill*. This algorithm leverages an iterative envelope-fitting approach using pseudo-convex bubbles to model broad fluorescence backgrounds without relying on polynomial or low-order parametric assumptions. It is particularly effective for biological samples, where low Raman-to-fluorescence signal ratios often present a challenge for standard methods.

— **Algorithm Performance Evaluation** : The performance of *BubbleFill* was quantita-

tively assessed against state-of-the-art baseline removal methods using both synthetic and empirical datasets. Results demonstrated that the algorithm consistently outperformed existing techniques in conditions characterized by high fluorescence and weak Raman signal contributions, validating its suitability for *in-vivo* tissue analysis.

— **Library Testing and Validation** : Beyond its integration with ORAS, ORPL was validated through its application to a wide range of Raman datasets acquired under varying experimental and clinical conditions. As detailed in Chapter 3, the datasets were acquired using three different Raman platforms. These included *in-vivo* brain tissue spectra (ODS Surgical system), *in-vivo* and *ex-vivo* prostate tissue (custom lab system), and paraffin-fixed prostate sections and dried saliva samples (Renishaw microscope). Together, the datasets comprised over 18,000 spectra, covering diverse sample types, preparation methods, and signal qualities. This broad validation confirms the robustness and versatility of ORPL for biomedical Raman spectroscopy.

— **Automatization of Processing Parameters** : Efforts were made to automatize the tuning of processing parameters through the development of the ASSI metric, which quantifies spectral sharpness and signal quality. However, results were inconsistent across and within datasets and did not yield a reliable method for parameter selection. As such, this sub-objective remains unmet, and the limitations encountered are discussed in the following section.

Objective 2 focused on the development of a modular Raman spectroscopy software platform enabling real-time acquisition, processing, and classification of *in-vivo* data. This objective was addressed through the design and deployment of the Lumed Raman System and the Open Raman Acquisition Software (ORAS), as presented in Chapter 4. The platform integrates hardware control, data processing, and diagnostic feedback into a cohesive interface built specifically for clinical and research use. The following sub-objectives contributed to the realization of this goal :

— **Modular Software Architecture** : ORAS was developed with a modular architecture that separates hardware control, acquisition logic, and user interface components into independent layers. Each hardware component (e.g., laser, camera, powermeter) is managed by a dedicated Python module, making it straightforward to extend the system to additional devices. This architecture simplifies maintenance and supports reuse in other research platforms.

— **Integration with Open-Source Processing Tools** : The ORPL library was fully integrated into ORAS, providing real-time preprocessing of acquired spectra. Spectral processing routines—including baseline removal, normalization, and smoothing—are

applied immediately after acquisition and before any classification or display, ensuring consistent and reproducible output across experiments.

— **Real-Time Processing and Classification** : ORAS includes a dynamic classification module that loads externally defined AI models for real-time tissue classification. A custom GUI widget was developed to display prediction outputs in a structured and interpretable format suited for clinical users. Specifications for model compatibility are defined in the ORAS documentation, enabling new models to be added by implementing a standard Python interface.

— **System Validation Using Phantom Models** and **Evaluation in Simulated Clinical Conditions** : These two sub-objectives were addressed through the deployment of the Lumed Raman System platform in real clinical environments. As detailed in Section 4.6.1, four Lumed Raman Systems running ORAS have been deployed across multiple hospital centres in Montreal. The systems have been actively used in several clinical studies, including *in-vivo* neurosurgical applications, surpassing the requirements of phantom testing and simulation. Data acquired using the platform has been published in peer-reviewed literature [5], further confirming its performance in realistic and high-stakes environments.

— **Usability and Deployment Efficiency** : The ORAS interface was designed for ease of use in time-constrained surgical workflows. Predefined acquisition profiles eliminate guesswork by encoding safe and effective parameters consistent with laser safety standards. Key acquisition routines—such as Automatic Exposure Control (AEC) and automatic accumulation optimization—further reduce operator burden. The entire measurement process, from system initialization to spectrum visualization and classification, can be completed in under 15 mouse clicks. As discussed in Section 4.7, these features contribute to the platform's usability, safety, and clinical relevance.

As a final note, the peer-reviewed publication associated with this work, presented in Chapter 3, was published in 2023 in the *Journal of Biomedical Optics*. As of April 2025, it has been cited 45 times according to Google Scholar and was recognized as JBO's most cited publication of 2023. This recognition was highlighted during Photonics West 2025 and later confirmed by the journal editors via correspondence to the authors.

5.2 Limits and Constraints

While this work achieved its primary objectives and produced a functional, open, and clinically validated Raman spectroscopy platform, several limitations and constraints remain.

A first limitation lies in the automatization of spectral processing parameters. Although the ASSI metric introduced in Chapter 3 showed promise as a quantitative measure of spectral quality, attempts to use it for automatic optimization of baseline removal and processing parameters were unsuccessful. The variability in signal characteristics across tissue types, acquisition environments, and instrumentation proved too complex to be robustly captured by a single quality metric. Specifically, because ASSI tends to favour spectra with a small number of sharp, narrow peaks—while being insensitive to absolute signal intensity due to normalization—optimal ASSI values were often obtained using extremely small bubble sizes in the *BubbleFill* algorithm. In these cases, baseline subtraction became overly aggressive, suppressing most of the true Raman signal and preserving only the very top of one or a few peaks. This behaviour occurred inconsistently across datasets and could not be systematically predicted or corrected. Similar outcomes were observed when alternative spectral quality metrics were employed. However, shifting to a different metric merely reintroduces the problem under a different guise—now the challenge becomes selecting which metric to trust. As such, fully automated parameter tuning in Raman processing remains an open problem requiring more adaptive or context-aware approaches.

Another limitation concerns the hardware dependencies of the Lumed Raman System. Although the platform is fully open-sourced and documented——with publicly available assembly guides, schematics, 3D-printed components, and compiled firmware for electronic subsystems——it still relies on several proprietary, off-the-shelf components. Chief among these are the EMVision spectrometer and associated optical hardware. While EMVision has been exceptionally collaborative and supportive of academic research, even going so far as to develop custom probe instruments and facilitate system-level modifications, the long-term availability and openness of these components cannot be guaranteed. A more restrictive corporate strategy could significantly hinder future development and replication efforts.

The Andor camera system presents an even more notable constraint. Operation of these detectors is entirely dependent on the proprietary Andor SDK, which is not freely distributed and must be licensed on a per-camera basis at significant monetary cost. The SDK imposes strict control over communication with the device, limiting flexibility and access for system integration. This contrasts sharply with other components such as the IPS laser or Thorlabs powermeter, which offer control via standard SCPI commands over serial interfaces. The broader issue reflected here is a structural one : many commercial hardware vendors maintain closed ecosystems that are incompatible with the principles of openness, adaptability, and low-barrier experimentation required in academic and early-stage translational research. This tension between openness in research and commercial viability poses an ongoing challenge for the sustainable development of clinically integrated optical systems.

Beyond system-level and implementation-specific constraints, Raman spectroscopy itself faces fundamental limitations that impact its clinical utility. One of the most significant remains the challenge posed by strong endogenous tissue fluorescence. While baseline correction algorithms—such as those implemented in ORPL—are standard for suppressing fluorescence background, they cannot recover meaningful Raman signal when fluorescence dominates the sensor's dynamic range and the Raman contribution falls below the noise floor. In such cases, even a theoretically optimal baseline removal algorithm yields spectra with insufficient signal-to-noise ratio for diagnostic use. This issue cannot be resolved through system optimization alone and often necessitates more complex and constrained alternatives such as Coherent Anti-Stokes Raman Spectroscopy (CARS), Surface-Enhanced Raman Spectroscopy (SERS), or photobleaching strategies. Unfortunately, the severity of fluorescence interference is highly tissue and application dependent, making it difficult to anticipate and account for in advance.

A second limitation arises from the inherently small sampling area of typical point-based Raman probes, which may be ill-suited for certain clinical tasks—particularly those involving large resection specimens such as breast or sarcoma tissues. The 3D scanning system described in Chapter 4 would offer a partial solution by enabling automated acquisition over larger surfaces; however, this approach is time-consuming and may not align well with intraoperative time constraints. In response to these limitations, current research in the field is increasingly oriented toward multispectral Raman imaging approaches. These systems trade spectral resolution for spatial coverage, enabling rapid interrogation of larger tissue areas. While promising for deployment at scale, such platforms depend on high-resolution point-based systems like the Lumed Raman System for their development—specifically for initial spectral exploration, band selection, and model training. This evolving landscape represents a critical area of ongoing research and is likely to shape the next generation of clinically viable Raman diagnostic technologies.

5.3 Recommendations

Building on the developments and findings presented in this thesis, several areas have been identified where focused efforts could lead to meaningful improvements. These recommendations are intended to guide future work aimed at enhancing the performance, usability, and clinical integration of Raman spectroscopy technologies, as well as expanding their applicability across broader biomedical research contexts.

A promising direction for future work is the integration of spectral processing parameter tuning into the training workflow of machine learning models. By treating processing parameters—such as baseline correction and normalization settings—as trainable hyperpa-

rameters within a machine learning pipeline, it may be possible to optimize them jointly with the classification model itself. This approach would effectively align preprocessing with downstream diagnostic objectives and could eliminate the remaining manual biases in Raman data analysis. Once the model is trained, the optimal parameters identified during training could then be fixed and applied for standard spectrum visualization and manual interpretation.

Another key recommendation concerns the calibration procedure of Raman systems, including the Lumed Raman System. Current calibration workflows rely on non-sterile reference materials such as Tylenol or nylon for wavenumber calibration (x-axis) and the NIST Raman standard for intensity response calibration (y-axis). While effective, these materials are incompatible with sterilization protocols, necessitating calibration after surgical procedures rather than before. Furthermore, the NIST standard is not widely available in all countries, presenting logistical barriers to reproducibility and routine use. Developing an alternative reference material that is both sterilizable and readily manufacturable would significantly improve the practicality of calibration in clinical settings. One promising approach involves the use of PDMS-based materials doped with broad-spectrum fluorophores such as nigrosin. These materials offer smooth fluorescence curves suitable for response calibration over the system's spectral range and could be produced in a sterilizable, disposable format, enabling more consistent and accessible calibration routines across sites.

Finally, one of the most impactful strategies for promoting the adoption of Raman spectroscopy in clinical and research settings is the development of accessible and well-structured training resources. Based on the experience accumulated throughout this project, it has become clear that fostering user understanding—rather than abstracting complexity away entirely—is often more effective in facilitating meaningful engagement with the technology. While simplifying interfaces and automating processes are valuable, they are no substitute for informed users who understand the underlying principles. Prioritizing the creation of clear, approachable documentation, tutorials, and usage guides may yield long-term benefits that surpass purely technical improvements. This emphasis on education and transparency aligns naturally with the open-access ethos that underpins the entire platform, and may prove essential in cultivating a broader and more informed user base.

5.4 Regulatory Considerations for Open-Source Medical Technologies in Canada

The integration of open-source software (OSS) into medical technologies—including those used for diagnostics, clinical decision support, and real-time analytics—presents both trans-

formative potential and substantial regulatory complexity. In Canada, the oversight of such technologies is governed by Health Canada, which classifies devices according to risk, from Class I (low risk) to Class IV (high risk) [148, 149]. All medical devices incorporating software must comply with standards such as ISO 13485 for quality management and IEC 62304 for software lifecycle processes. These standards mandate traceability, validation, and post-market surveillance—all of which can pose unique challenges for software released under OSS licenses.

A central legal conflict arises from the contrasting legal responsibilities in the OSS and regulatory domains. Most open-source licenses include clauses that explicitly waive liability for developers, whereas Health Canada requires a clearly identifiable manufacturer to assume full legal responsibility for device performance and safety [102]. This contradiction presents a currently unresolved regulatory tension : OSS encourages decentralized, community-based development free from liability, while Health Canada mandates centralized, accountable oversight. At present, no frameworks exist that fully reconcile these diverging principles.

Economic considerations further complicate adoption. Medical technologies are expensive to develop and often require years of research, regulatory engagement, and clinical validation. Open-sourcing software components may appear to diminish competitive advantage, reducing investor appeal due to perceived loss of intellectual property protection. However, OSS can provide strategic benefits through first-mover advantage, accelerated development cycles, and collaborative infrastructure. These factors are increasingly significant in domains such as diagnostic software infrastructures and clinical data pipelines, where interoperability and transparency are paramount [103, 104].

In the Canadian context, health technology assessment (HTA) also influences adoption decisions. HTA frameworks increasingly consider not only clinical efficacy but also economic and organizational impact. As Polisena et al. note, this multi-criteria approach may favour open-source technologies in publicly funded healthcare systems, where affordability, transparency, and adaptability are desirable [150].

The tension between openness and regulatory rigour is especially pronounced in the context of artificial intelligence (AI), which is rapidly becoming embedded in medical diagnostics, image analysis, and decision-support tools. As regulators begin to scrutinize AI systems more intensively, transparency and auditability are emerging as foundational requirements. The Open-Source Initiative's *Open Source AI Definition — 1.0* reflects this shift, proposing that AI systems be considered "open" only if they grant four essential freedoms : to use, study, modify, and share the system and its components. Importantly, it emphasizes the need for access to source code, model parameters, and detailed information about the training

data—such as provenance, selection criteria, and processing methods—even if the data itself is not fully redistributed [151].

This approach acknowledges that some training data may be legally, ethically, or practically unsharable, and explicitly avoids requiring public release of datasets. Instead, it prioritizes reproducibility and informed modification by ensuring transparency about how the data was acquired and used. In AI-enabled medical applications, where safety risks may stem from biased outputs or opaque data handling, such structured openness offers a pragmatic pathway for building regulatory trust. Open systems that provide verifiable insight into training processes and system behaviour may ultimately be better positioned to satisfy emerging regulatory expectations in clinical environments.

In summary, while the OSS paradigm challenges traditional regulatory and commercial assumptions, it also offers significant advantages in transparency, scalability, and public accountability. The development of new legal instruments, risk-sharing models, and hybrid governance strategies will be essential to unlock the full potential of OSS in regulated medical environments. As both technological and regulatory landscapes evolve, Canada has the opportunity to lead in defining frameworks that support open, safe, and equitable innovation in medical devices and softwares.

5.5 Final Remarks

This thesis represents an effort not only to advance the technical and clinical applicability of Raman spectroscopy, but also to reshape how such technologies are developed, shared, and adopted. By grounding this work in the principles of openness, modularity, and usability, it contributes to a growing movement toward more transparent and collaborative approaches in biomedical engineering. The systems, tools, and methods developed here were designed to be used, modified, and improved by others, ensuring that their value extends beyond a single PhD student, research group or institution. While challenges will always remain, the results presented in this work suggest that Raman spectroscopy, when developed with intention and openness, can move closer to becoming a reliable and impactful tool in the healthcare landscape. Continued progress will depend not only on engineering breakthroughs, but also on sustained efforts to share knowledge, train users, and support an engaged community of researchers and clinicians who can drive the technology forward.

5.6 Other Scientific Contributions

In addition to the work presented in this thesis, I have contributed significantly to a range of collaborative research efforts in the field of biomedical optics, with a particular focus on Raman spectroscopy, machine learning, and clinical translation. These contributions span multiple areas of investigation, including system development, software architecture, experimental protocol design, data acquisition and analysis, and manuscript preparation. Many of these projects were conducted in collaboration with multidisciplinary teams of engineers, physicists, clinicians, and biologists, reflecting the highly integrative nature of translational medical research.

Table 5.1 summarizes the peer-reviewed journal articles to which I have contributed throughout the duration of my doctoral studies.

TABLE 5.1 – Peer-reviewed publications with contributions by G. Sheehy

Year	Author Pos.	Title	Journal	Ref.
2025	5th	An experimental method to assess depth sensing limits of inelastic scattering measurements using spatial-offset Raman spectroscopy imaging	Journal of Biomedical Optics	in submission
2025	3rd	Development and preclinical evaluation of an endonasal Raman spectroscopy probe for transsphenoidal pituitary adenoma surgery	Journal of Biomedical Optics	[5]
2024	7th	Quantitative assessment of the generalizability of a brain tumor Raman spectroscopy machine learning model to various tumor types including astrocytoma and oligodendroglioma	Journal of Biomedical Optics	[152]
2024	4th	Liquid saliva-based Raman spectroscopy device with on-board machine learning detects COVID-19 infection in real-time	Analyst	[153]

Continued on next page

Table 5.1 – continued from previous page

Year	Author Pos.	Title	Journal	Ref.
2024	4th	In Situ Brain Tumor Detection Using the Raman Spectroscopy Sentry System — Results of a Multicenter Study	Scientific Reports	[56]
2024	3rd	Spectral effects and enhancement quantification in healthy human saliva with surface-enhanced Raman spectroscopy using silver nanopillar substrates	Lasers in Surgery and Medicine	[154]
2023	4th	In situ Raman spectroscopy and machine learning unveil biomolecular alterations in breast cancer	Journal of Biomedical Optics	[48]
2023	1st	Open-sourced Raman spectroscopy data processing package implementing a baseline removal algorithm validated from multiple datasets acquired in human tissue and biofluids	Journal of Biomedical Optics	[147]
2022	4th	Image-guided Raman spectroscopy navigation system to improve transperineal prostate cancer detection. Part 2 : Machine learning model combining spectral with radiomics features improves in situ tumor-targeting accuracy	Journal of Biomedical Optics	[52]
2022	4th	Image-guided Raman spectroscopy navigation system to improve transperineal prostate cancer detection. Part 1 : Overview of Raman spectroscopy fiber-optics system and in-situ tissue characterization	Journal of Biomedical Optics	[51]

Continued on next page

Table 5.1 – continued from previous page

Year	Author Pos.	Title	Journal	Ref.
2021	5th	Saliva-based detection of COVID-19 infection in a real-world setting using reagent-free Raman spectroscopy and machine learning	Journal of Biomedical Optics	[3]
2021	5th	Multispectral label-free Raman spectroscopy can detect ovarian and endometrial cancer with high accuracy	Journal of Biophotonics	[155]
2021	3rd	Data consistency and classification model transferability across biomedical Raman spectroscopy systems	Translational Biophotonics	[135]
2020	4th	Pre-clinical evaluation of an image-guided in-situ Raman spectroscopy navigation system for targeted prostate cancer interventions	International Journal of Computer Assisted Radiology and Surgery	[156]
2020	4th	Quantitative spectral quality assessment technique validated using intraoperative in vivo Raman spectroscopy measurements	Journal of Biomedical Optics	[137]
2020	3rd	Experimental validation of a spectroscopic Monte Carlo light transport simulation technique and Raman scattering depth sensing analysis in biological tissue	Journal of Biomedical Optics	[157]

Continued on next page

Table 5.1 – continued from previous page

Year	Author Pos.	Title	Journal	Ref.
2019	4th	Wide-Field Spatial Frequency Domain Imaging, Diffuse Reflectance and Endogenous Fluorescence Spectroscopy System for Quantitative Tissue Biomarkers in Radical Prostatectomy Specimens	Biomedical Optics Express	[158]

REFERENCES

[1] Gregory W. Auner, S. Kiran Koya, Changhe Huang, Brandy Broadbent, Micaela Trexler, Zachary Auner, Angela Elias, Katlyn Curtin Mehne, and Michelle A. Brusatori. Applications of Raman spectroscopy in cancer diagnosis. *Cancer and Metastasis Reviews*, 37(4) :691–717, December 2018. ISSN 0167-7659, 1573-7233. doi : 10.1007/s10555-018-9770-9.

[2] E. B. Hanlon, R. Manoharan, T.-W. Koo, K. E. Shafer, J. T. Motz, M. Fitzmaurice, J. R. Kramer, I. Itzkan, R. R. Dasari, and M. S. Feld. Prospects for in vivo Raman spectroscopy. *Physics in Medicine & Biology*, 45(2) :R1, February 2000. ISSN 0031-9155. doi : 10.1088/0031-9155/45/2/201.

[3] Katherine Ember, François Daoust, Myriam Mahfoud, Frédérick Dallaire, Esmat Zamani Ahmad, Trang Tran, Arthur Plante, Mame-Kany Diop, Tien Nguyen, Amélie St-Georges-Robillard, Nassim Ksantini, Julie Lanthier, Antoine Filiatrault, Guillaume Sheehy, Gabriel Beaudoin, Caroline Quach, Dominique Trudel, and Frédéric Leblond. Saliva-based detection of COVID-19 infection in a real-world setting using reagent-free Raman spectroscopy and machine learning. *Journal of Biomedical Optics*, 27(02), February 2022. ISSN 1083-3668. doi : 10.1117/1.JBO.27.2.025002.

[4] Eric Todd Marple, Kirk David Urmey, and John Meckert. Fiber optic probe having components and features that improve performance, October 2022.

[5] Victor Blanquez-Yeste, Félix Janelle, Trang Tran, Katherine Ember, Guillaume Sheehy, Frédérick Dallaire, Eric Marple, Kirk Urmey, Moujahed Labidi, and Frédéric Leblond. Development and preclinical evaluation of an endonasal Raman spectroscopy probe for transsphenoidal pituitary adenoma surgery. *Journal of Biomedical Optics*, 30(3) : 035004, March 2025. ISSN 1083-3668, 1560-2281. doi : 10.1117/1.JBO.30.3.035004.

[6] Eric Todd Marple, Kirk David Urmey, and John Meckert. Spectrometers having a variable focus lens with a single air gap, March 2022.

[7] Kenny Kong, Catherine Kendall, Nicholas Stone, and Ioan Notingher. Raman spectroscopy for medical diagnostics — From in-vitro biofluid assays to in-vivo cancer detection. *Advanced Drug Delivery Reviews*, 89 :121–134, July 2015. ISSN 0169409X. doi : 10.1016/j.addr.2015.03.009.

[8] Miguel Ghebré Ramírez-Elías, Francisco Javier González, Miguel Ghebré Ramírez-Elías, and Francisco Javier González. Raman Spectroscopy for In Vivo Medical Diagnosis. In *Raman Spectroscopy*. IntechOpen, February 2018. ISBN 978-1-78923-001-7.

doi : 10.5772/intechopen.72933.

[9] Michael Jermyn, Kelvin Mok, Jeanne Mercier, Joannie Desroches, Julien Pichette, Karl Saint-Arnaud, Liane Bernstein, Marie-Christine Guiot, Kevin Petrecca, and Frederic Leblond. Intraoperative brain cancer detection with Raman spectroscopy in humans. *Science Translational Medicine*, 7(274) :274ra19–274ra19, February 2015. ISSN 1946-6234, 1946-6242. doi : 10.1126/scitranslmed.aaa2384.

[10] Lauren A. Austin, Sam Osseiran, and Conor L. Evans. Raman technologies in cancer diagnostics. *The Analyst*, 141(2) :476–503, 2016. ISSN 0003-2654, 1364-5528. doi : 10.1039/C5AN01786F.

[11] Inês P. Santos, Elisa M. Barroso, Tom C. Bakker Schut, Peter J. Caspers, Cornelia G. F. van Lanschot, Da-Hye Choi, Martine F. van der Kamp, Roeland W. H. Smits, Remco van Doorn, Rob M. Verdijk, Vincent Noordhoek Hegt, Jan H. von der Thüsen, Carolien H. M. van Deurzen, Linetta B. Koppert, Geert J. L. H. van Leenders, Patricia C. Ewing-Graham, Helena C. van Doorn, Clemens M. F. Dirven, Martijn B. Busstra, Jose Hardillo, Aniel Sewnaik, Ivo ten Hove, Hetty Mast, Dominiek A. Monserez, Cees Meeuwis, Tamar Nijsten, Eppo B. Wolvius, Robert J. Baatenburg de Jong, Gerwin J. Puppels, and Senada Koljenović. Raman spectroscopy for cancer detection and cancer surgery guidance : Translation to the clinics. *The Analyst*, 142(17) :3025–3047, 2017. ISSN 0003-2654, 1364-5528. doi : 10.1039/C7AN00957G.

[12] Yulia Khristoforova, Lyudmila Bratchenko, and Ivan Bratchenko. Raman-Based Techniques in Medical Applications for Diagnostic Tasks : A Review. *International Journal of Molecular Sciences*, 24(21) :15605, January 2023. ISSN 1422-0067. doi : 10.3390/ijms242115605.

[13] Yaping Qi, Esther Xinyi Chen, Dan Hu, Ying Yang, Zhenping Wu, Ming Zheng, Mohammad A. Sadi, Yucheng Jiang, Kang Zhang, Zi Chen, and Yong P. Chen. Applications of Raman spectroscopy in clinical medicine. *Food Frontiers*, 5(2) :392–419, 2024. ISSN 2643-8429. doi : 10.1002/fft2.335.

[14] Lukas Simon Kriem, Kevin Wright, Renzo Alberto Ccahuana-Vasquez, and Steffen Rupp. Confocal Raman microscopy to identify bacteria in oral subgingival biofilm models. *PLOS ONE*, 15(5) :e0232912, May 2020. ISSN 1932-6203. doi : 10.1371/journal.pone.0232912.

[15] Leonardo M. Moreira, Landulfo Silveira Jr., Fábio V. Santos, Juliana P. Lyon, Rick Rocha, Renato A. Zângaro, Antonio Balbin Villaverde, and Marcos T. T. Pacheco. Raman spectroscopy : A powerful technique for biochemical analysis and diagnosis. *Journal of Spectroscopy*, 22(1) :942758, 2008. ISSN 2314-4939. doi : 10.3233/SPE-2008-0326.

[16] Alois Bonifacio, Silvia Cervo, and Valter Sergo. Label-free surface-enhanced Raman spectroscopy of biofluids : Fundamental aspects and diagnostic applications. *Analytical and Bioanalytical Chemistry*, 407(27) :8265–8277, November 2015. ISSN 1618-2642, 1618-2650. doi : 10.1007/s00216-015-8697-z.

[17] Corina D. Hosu, Vlad Moisoiu, Andrei Stefancu, Elisabeta Antonescu, Loredana F. Leopold, Nicolae Leopold, and Daniela Fodor. Raman spectroscopy applications in rheumatology. *Lasers in Medical Science*, 34(4) :827–834, June 2019. ISSN 0268-8921, 1435-604X. doi : 10.1007/s10103-019-02719-2.

[18] Noureen Siraj, Bwambok , David K., Brady , Pamela Nicole, Taylor , Megan, Baker , Gary A., Bashiru , Mujeebat, Macchi , Samantha, Jalihal , Amanda, Denmark , Iris, Le , Thuy, Elzey , Brianda, Pollard , David A., and Sayo O. and Fakayode. Raman spectroscopy and multivariate regression analysis in biomedical research, medical diagnosis, and clinical analysis. *Applied Spectroscopy Reviews*, 56(8-10) :615–672, November 2021. ISSN 0570-4928. doi : 10.1080/05704928.2021.1913744.

[19] Émile Lemoine, Frédérick Dallaire, Rajeev Yadav, Rajeev Agarwal, Samuel Kadoury, Dominique Trudel, Marie-Christine Guiot, Kevin Petrecca, and Frédéric Leblond. Feature engineering applied to intraoperative in vivo Raman spectroscopy sheds light on molecular processes in brain cancer : A retrospective study of 65 patients. *Analyst*, 144 (22) :6517–6532, November 2019. ISSN 1364-5528. doi : 10.1039/C9AN01144G.

[20] Andrew M. Fales, Ilko K. Ilev, and T. Joshua Pfefer. Evaluation of standardized performance test methods for biomedical Raman spectroscopy. *Journal of Biomedical Optics*, 27(07), October 2021. ISSN 1083-3668. doi : 10.1117/1.JBO.27.7.074705.

[21] Walter Stummer, Uwe Pichlmeier, Thomas Meinel, Otmar Dieter Wiestler, Friedhelm Zanella, Hans-Jürgen Reulen, and ALA-Glioma Study Group. Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma : A randomised controlled multicentre phase III trial. *The Lancet. Oncology*, 7(5) :392–401, May 2006. ISSN 1470-2045. doi : 10.1016/S1470-2045(06)70665-9.

[22] Constantinos G. Hadjipanayis, Walter Stummer, and Jason P. Sheehan. 5-ALA fluorescence-guided surgery of CNS tumors. *Journal of Neuro-Oncology*, 141(3) :477–478, February 2019. ISSN 1573-7373. doi : 10.1007/s11060-019-03109-y.

[23] Friso B. Achterberg, Babs G. Sibinga Mulder, Ruben P. J. Meijer, Bert A. Bonsing, Henk H. Hartgrink, J. Sven D. Mieog, Aimen Zlitni, Seung-min Park, Arantza Farina Sarasqueta, Alexander L. Vahrmeijer, and Rutger-Jan Swijnenburg. Real-time surgical margin assessment using ICG-fluorescence during laparoscopic and robot-assisted resections of colorectal liver metastases. *Annals of Translational Medicine*, 8(21) : 1448–1448, November 2020. ISSN 23055839, 23055847. doi : 10.21037/atm-20-1999.

[24] Benjamin Weixler, Leonard A. Lobbes, Luis Scheiner, Johannes C. Lauscher, Sebastian M. Staubli, Markus Zuber, and Dimitri A. Raptis. The Value of Indocyanine Green Image-Guided Surgery in Patients with Primary Liver Tumors and Liver Metastases. *Life*, 13(6) :1290, May 2023. ISSN 2075-1729. doi : 10.3390/life13061290.

[25] Giri Krishnan, Nynke S. Van Den Berg, Naoki Nishio, Shrey Kapoor, Jaqueline Pei, Laura Freeman, Yu-Jin Lee, Quan Zhou, Stan Van Keulen, Shayan Farkurnejad, James Condon, Fred M. Baik, Brock A. Martin, and Eben L. Rosenthal. Fluorescent Molecular Imaging Can Improve Intraoperative Sentinel Margin Detection in Oral Squamous Cell Carcinoma. *Journal of Nuclear Medicine*, 63(8) :1162–1168, August 2022. ISSN 0161-5505, 2159-662X. doi : 10.2967/jnumed.121.262235.

[26] F.T. Nguyen, A.M. Zysk, E.J. Chaney, S.G. Adie, J.G. Kotynek, U.J. Oliphant, F.J. Bellafiore, K.M. Rowland, P.A. Johnson, and S.A. Boppart. Optical Coherence Tomography : The Intraoperative Assessment of Lymph Nodes in Breast Cancer. *IEEE Engineering in Medicine and Biology Magazine*, 29(2) :63–70, March 2010. ISSN 0739-5175. doi : 10.1109/MEMB.2009.935722.

[27] Adam M. Zysk, Kai Chen, Edward Gabrielson, Lorraine Tafra, Evelyn A. May Gonzalez, Joseph K. Canner, Eric B. Schneider, Andrew J. Cittadine, P. Scott Carney, Stephen A. Boppart, Kimiko Tsuchiya, Kristen Sawyer, and Lisa K. Jacobs. Intraoperative Assessment of Final Margins with a Handheld Optical Imaging Probe During Breast-Conserving Surgery May Reduce the Reoperation Rate : Results of a Multicenter Study. *Annals of Surgical Oncology*, 22(10) :3356–3362, October 2015. ISSN 1068-9265, 1534-4681. doi : 10.1245/s10434-015-4665-2.

[28] Shuwei Zhang, Bin Yang, Houpu Yang, Jin Zhao, Yuanyuan Zhang, Yuanxu Gao, Olivia Monteiro, Kang Zhang, Bo Liu, and Shu Wang. Potential rapid intraoperative cancer diagnosis using dynamic full-field optical coherence tomography and deep learning : A prospective cohort study in breast cancer patients. *Science Bulletin*, 69(11) :1748–1756, June 2024. ISSN 2095-9273. doi : 10.1016/j.scib.2024.03.061.

[29] Júlia Balog, László Sasi-Szabó, James Kinross, Matthew R. Lewis, Laura J. Muirhead, Kirill Veselkov, Reza Mirnezami, Balázs Dezső, László Damjanovich, Ara Darzi, Jeremy K. Nicholson, and Zoltán Takáts. Intraoperative Tissue Identification Using Rapid Evaporative Ionization Mass Spectrometry. *Science Translational Medicine*, 5(194), July 2013. ISSN 1946-6234, 1946-6242. doi : 10.1126/scitranslmed.3005623.

[30] Mary E. King, Jialing Zhang, John Q. Lin, Kyana Y. Garza, Rachel J. DeHoog, Clara L. Feider, Alena Bensussan, Marta Sans, Anna Krieger, Sunil Badal, Michael F. Keating, Spencer Woody, Sadhna Dhingra, Wendong Yu, Christopher Pirko, Kirtan A. Brahmbhatt, George Van Buren, William E. Fisher, James Suliburk, and Livia S. Eberlin. Ra-

pid diagnosis and tumor margin assessment during pancreatic cancer surgery with the MasSpec Pen technology. *Proceedings of the National Academy of Sciences*, 118(28) : e2104411118, July 2021. ISSN 0027-8424, 1091-6490. doi : 10.1073/pnas.2104411118.

[31] Francesco Prada, Luca Mattei, Massimiliano Del Bene, Luca Aiani, Marco Saini, Cecilia Casali, Assunta Filippini, Federico Giuseppe Legnani, Alessandro Perin, Andrea Saladino, Ignazio Gaspare Vetrano, Luigi Solbiati, Alberto Martegani, and Francesco DiMeco. Intraoperative cerebral glioma characterization with contrast enhanced ultrasound. *BioMed Research International*, 2014 :484261, 2014. ISSN 2314-6141. doi : 10.1155/2014/484261.

[32] Ling-Gang Cheng, Wen He, Hong-Xia Zhang, Qian Song, Bin Ning, Hui-Zhan Li, Yan He, and Song Lin. Intraoperative Contrast Enhanced Ultrasound Evaluates the Grade of Glioma. *BioMed Research International*, 2016 :2643862, 2016. ISSN 2314-6141. doi : 10.1155/2016/2643862.

[33] Christian Senft, Andrea Bink, Kea Franz, Hartmut Vatter, Thomas Gasser, and Volker Seifert. Intraoperative MRI guidance and extent of resection in glioma surgery : A randomised, controlled trial. *The Lancet. Oncology*, 12(11) :997–1003, October 2011. ISSN 1474-5488. doi : 10.1016/S1470-2045(11)70196-6.

[34] Michael D Jenkinson, Damiano Giuseppe Barone, Andrew Bryant, Luke Vale, Helen Bulbeck, Theresa A Lawrie, Michael G Hart, and Colin Watts. Intraoperative imaging technology to maximise extent of resection for glioma. *Cochrane Database of Systematic Reviews*, 2021(5), January 2018. ISSN 14651858. doi : 10.1002/14651858.CD012788.pub2.

[35] Barbara L. Smith, Conor R. Lanahan, Michelle C. Specht, Bridget N. Kelly, Carson Brown, David B. Strasfeld, Jorge M. Ferrer, Upahvan Rai, Rong Tang, Travis Rice-Stitt, Anna Biernacka, Elena F. Brachtel, and Michele A. Gadd. Feasibility Study of a Novel Protease-Activated Fluorescent Imaging System for Real-Time, Intraoperative Detection of Residual Breast Cancer in Breast Conserving Surgery. *Annals of Surgical Oncology*, 27(6) :1854–1861, June 2020. ISSN 1534-4681. doi : 10.1245/s10434-019-08158-1.

[36] Stan Van Keulen, Naoki Nishio, Shayan Fakurnejad, Nynke S. Van Den Berg, Guolan Lu, Andrew Birkeland, Brock A. Martin, Tymour Forouzanfar, Dimitrios A Colevas, and Eben L. Rosenthal. Intraoperative Tumor Assessment Using Real-Time Molecular Imaging in Head and Neck Cancer Patients. *Journal of the American College of Surgeons*, 229(6) :560–567e1, December 2019. ISSN 1072-7515. doi : 10.1016/j.jamcollsurg.2019.09.007.

[37] Dhurka Shanthakumar, Maria Leiloglou, Colm Kelliher, Ara Darzi, Daniel S. Elson, and Daniel R. Leff. A Comparison of Spectroscopy and Imaging Techniques Utilizing

Spectrally Resolved Diffusely Reflected Light for Intraoperative Margin Assessment in Breast-Conserving Surgery : A Systematic Review and Meta-Analysis. *Cancers*, 15 (11) :2884, May 2023. ISSN 2072-6694. doi : 10.3390/cancers15112884.

[38] Jiongru Pan, Han Deng, Shiqi Hu, Chengwan Xia, Yongfeng Chen, Jianquan Wang, and Yuxin Wang. Real-time surveillance of surgical margins via ICG-based near-infrared fluorescence imaging in patients with OSCC. *World Journal of Surgical Oncology*, 18 (1) :96, December 2020. ISSN 1477-7819. doi : 10.1186/s12957-020-01874-z.

[39] Florin-Catalin Pop, Isabelle Veys, Sophie Vankerckhove, Romain Barbieux, Marie Chintinne, Michel Moreau, Vincent Donckier, Denis Larsimont, Pierre Bourgeois, and Gabriel Liberale. Absence of residual fluorescence in the surgical bed at near-infrared fluorescence imaging predicts negative margins at final pathology in patients treated with breast-conserving surgery for breast cancer. *European Journal of Surgical Oncology : The Journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology*, 47(2) :269–275, February 2021. ISSN 1532-2157. doi : 10.1016/j.ejso.2020.09.036.

[40] Jennifer Eschbacher, Nikolay L. Martirosyan, Peter Nakaji, Nader Sanai, Mark C. Preul, Kris A. Smith, Stephen W. Coons, and Robert F. Spetzler. In vivo intraoperative confocal microscopy for real-time histopathological imaging of brain tumors : Clinical article. *Journal of Neurosurgery*, 116(4) :854–860, April 2012. ISSN 0022-3085, 1933-0693. doi : 10.3171/2011.12.JNS11696.

[41] M. Sam Eljamel and Syed Osama Mahboob. The effectiveness and cost-effectiveness of intraoperative imaging in high-grade glioma resection ; a comparative review of intraoperative ALA, fluorescein, ultrasound and MRI. *Photodiagnosis and Photodynamic Therapy*, 16 :35–43, December 2016. ISSN 15721000. doi : 10.1016/j.pdpdt.2016.07.012.

[42] Marc Thill. MarginProbe : Intraoperative margin assessment during breast conserving surgery by using radiofrequency spectroscopy. *Expert Review of Medical Devices*, 10 (3) :301–315, May 2013. ISSN 1745-2422. doi : 10.1586/erd.13.5.

[43] Stan Van Keulen, Naoki Nishio, Shayan Fakurnejad, Andrew Birkeland, Brock A. Martin, Guolan Lu, Quan Zhou, Stefania U. Chirita, Tymour Forouzanfar, A. Dimitrios Colevas, Nynke S. Van Den Berg, and Eben L. Rosenthal. The Clinical Application of Fluorescence-Guided Surgery in Head and Neck Cancer. *Journal of Nuclear Medicine*, 60(6) :758–763, June 2019. ISSN 0161-5505, 2159-662X. doi : 10.2967/jnumed.118.222810.

[44] Kurtis Young, Enze Ma, Sameer Kejriwal, Torbjoern Nielsen, Sukhkaran S. Aulakh, and Andrew C. Birkeland. Intraoperative In Vivo Imaging Modalities in Head and

Neck Cancer Surgical Margin Delineation : A Systematic Review. *Cancers*, 14(14) : 3416, July 2022. ISSN 2072-6694. doi : 10.3390/cancers14143416.

[45] Abigail S. Haka, Karen E. Shafer-Peltier, Maryann Fitzmaurice, Joseph Crowe, Ramachandra R. Dasari, and Michael S. Feld. Diagnosing breast cancer by using Raman spectroscopy. *Proceedings of the National Academy of Sciences*, 102(35) :12371–12376, August 2005. doi : 10.1073/pnas.0501390102.

[46] Ashley M. Laughney, Venkataramanan Krishnaswamy, Elizabeth J. Rizzo, Mary C. Schwab, Richard J. Barth, Brian W. Pogue, Keith D. Paulsen, and Wendy A. Wells. Scatter spectroscopic imaging distinguishes between breast pathologies in tissues relevant to surgical margin assessment. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 18(22) :6315–6325, November 2012. ISSN 1078-0432. doi : 10.1158/1078-0432.CCR-12-0136.

[47] Jakub Surmacki, Jacek Musial, Radzislaw Kordek, and Halina Abramczyk. Raman imaging at biological interfaces : Applications in breast cancer diagnosis. *Molecular Cancer*, 12(1) :1–12, December 2013. ISSN 1476-4598. doi : 10.1186/1476-4598-12-48.

[48] Sandryne David, Trang Tran, Frédérick Dallaire, Guillaume Sheehy, Feryel Azzi, Dominique Trudel, Francine Tremblay, Atilla Omeroglu, Frédéric Leblond, and Sarkis Meterissian. In situ Raman spectroscopy and machine learning unveil biomolecular alterations in invasive breast cancer. *Journal of Biomedical Optics*, 28(3) :036009, March 2023. ISSN 1560-2281. doi : 10.1117/1.JBO.28.3.036009.

[49] Nicholas Stone, Catherine Kendall, Jenny Smith, Paul Crow, and Hugh Barr. Raman spectroscopy for identification of epithelial cancers. *Faraday Discussions*, 126(0) :141–157, December 2004. ISSN 1364-5498. doi : 10.1039/B304992B.

[50] Wooje Lee, Aufried T.M. Lenferink, Cees Otto, and Herman L. Offerhaus. Classifying Raman spectra of extracellular vesicles based on convolutional neural networks for prostate cancer detection. *Journal of Raman Spectroscopy*, 51(2) :293–300, 2020. ISSN 1097-4555. doi : 10.1002/jrs.5770.

[51] Fabien Picot, Roozbeh Shams, Frédérick Dallaire, Guillaume Sheehy, Tran Trang, David Grajales, Mirela Birlea, Dominique Trudel, Cynthia Ménard, Samuel Kadoury, and Frédéric Leblond. Image-guided Raman spectroscopy navigation system to improve transperineal prostate cancer detection. Part 1 : Raman spectroscopy fiber-optics system and in situ tissue characterization. *Journal of Biomedical Optics*, 27(9) :095003, September 2022. ISSN 1560-2281. doi : 10.1117/1.JBO.27.9.095003.

[52] David Orlando Grajales Lopera, Fabien Picot, Roozbeh Shams, Frédérick Dallaire, Guillaume Sheehy, Stephanie Alley, Maroie Barkati, Guila Delouya, Jean-Francois Carrier, Mirela Birlea, Dominique Trudel, Frédéric Leblond, Cynthia Ménard, and Samuel

Kadoury. Image-guided Raman spectroscopy navigation system to improve transperineal prostate cancer detection. Part 2 : In-vivo tumor-targeting using a classification model combining spectral and MRI-radiomics features. *Journal of Biomedical Optics*, 27(9) :095004, September 2022. ISSN 1560-2281. doi : 10.1117/1.JBO.27.9.095004.

[53] Maria Giovanna Lizio, Radu Boitor, and Ioan Notingher. Selective-sampling Raman imaging techniques for ex vivo assessment of surgical margins in cancer surgery. *Analyst*, 146(12) :3799–3809, June 2021. ISSN 1364-5528. doi : 10.1039/D1AN00296A.

[54] Xu Feng, Austin J. Moy, Hieu T. M. Nguyen, Yao Zhang, Jason Zhang, Matthew C. Fox, Katherine R. Sebastian, Jason S. Reichenberg, Mia K. Markey, and James W. Tunnell. Raman biophysical markers in skin cancer diagnosis. *Journal of Biomedical Optics*, 23(5) :1–10, May 2018. ISSN 1560-2281. doi : 10.1117/1.JBO.23.5.057002.

[55] Joannie Desroches, Michael Jermyn, Michael Pinto, Fabien Picot, Marie-Andrée Tremblay, Sami Obaid, Eric Marple, Kirk Urmey, Dominique Trudel, Gilles Soulez, Marie-Christine Guiot, Brian C. Wilson, Kevin Petrecca, and Frédéric Leblond. A new method using Raman spectroscopy for in vivo targeted brain cancer tissue biopsy. *Scientific Reports*, 8(1) :1792, January 2018. ISSN 2045-2322. doi : 10.1038/s41598-018-20233-3.

[56] Katherine Ember, Frédérick Dallaire, Arthur Plante, Guillaume Sheehy, Marie-Christine Guiot, Rajeev Agarwal, Rajeev Yadav, Alice Douet, Juliette Selb, Jean Philippe Tremblay, Alex Dupuis, Eric Marple, Kirk Urmey, Caroline Rizea, Armand Harb, Lily McCarthy, Alexander Schupper, Melissa Umphlett, Nadejda Tsankova, Frédéric Leblond, Constantinos Hadjipanayis, and Kevin Petrecca. In situ brain tumor detection using a Raman spectroscopy system—results of a multicenter study. *Scientific Reports*, 14(1) :13309, June 2024. ISSN 2045-2322. doi : 10.1038/s41598-024-62543-9.

[57] Steven L. Jacques and Brian W. Pogue. Tutorial on diffuse light transport. *Journal of Biomedical Optics*, 13(4) :041302, 2008. ISSN 10833668. doi : 10.1117/1.2967535.

[58] Ramasamy Manoharan, Yang Wang, and Michael S. Feld. Histochemical analysis of biological tissues using Raman spectroscopy. *Spectrochimica Acta Part A : Molecular and Biomolecular Spectroscopy*, 52(2) :215–249, February 1996. ISSN 1386-1425. doi : 10.1016/0584-8539(95)01573-6.

[59] Peter L. Choyke and Hisataka Kobayashi. Medical Uses of Fluorescence Imaging : Bringing Disease to Light. *IEEE Journal of Selected Topics in Quantum Electronics*, 18(3) :1140–1146, May 2012. ISSN 1558-4542. doi : 10.1109/JSTQE.2011.2164900.

[60] Evgenii Belykh, Nikolay L. Martirosyan, Kaan Yagmurlu, Eric J. Miller, Jennifer M. Eschbacher, Mohammadhassan Izadyyazdanabadi, Liudmila A. Bardonova, Vadim A. Byvaltsev, Peter Nakaji, and Mark C. Preul. Intraoperative Fluorescence Imaging for

Personalized Brain Tumor Resection : Current State and Future Directions. *Frontiers in Surgery*, 3, October 2016. ISSN 2296-875X. doi : 10.3389/fsurg.2016.00055.

[61] Steve S. Cho, Ryan Salinas, and John Y. K. Lee. Indocyanine-Green for Fluorescence-Guided Surgery of Brain Tumors : Evidence, Techniques, and Practical Experience. *Frontiers in Surgery*, 6, March 2019. ISSN 2296-875X. doi : 10.3389/fsurg.2019.00011.

[62] Thinzar M. Lwin, Robert M. Hoffman, and Michael Bouvet. Fluorescence-guided hepatobiliary surgery with long and short wavelength fluorophores. *Hepatobiliary Surgery and Nutrition*, 9(5) :61539–61639, October 2020. ISSN 2304-389X, 2304-3881. doi : 10.21037/hbsn.2019.09.13.

[63] Alexander J. Schupper, Manasa Rao, Nicki Mohammadi, Rebecca Baron, John Y. K. Lee, Francesco Acerbi, and Constantinos G. Hadjipanayis. Fluorescence-Guided Surgery : A Review on Timing and Use in Brain Tumor Surgery. *Frontiers in Neurology*, 12, June 2021. ISSN 1664-2295. doi : 10.3389/fneur.2021.682151.

[64] Dimitris Gorpas and Laura Marcu. Fluorescence Lifetime Spectroscopy and Imaging Techniques in Medical Applications. In Malini Olivo and U. S. Dinish, editors, *Frontiers in Biophotonics for Translational Medicine : In the Celebration of Year of Light (2015)*, pages 1–46. Springer, Singapore, 2016. ISBN 978-981-287-627-0. doi : 10.1007/978-981-287-627-0_1.

[65] Craig M. Gardner, Steven L. Jacques, and Ashley J. Welch. Fluorescence spectroscopy of tissue : Recovery of intrinsic fluorescence from measured fluorescence. *Applied Optics*, 35(10) :1780–1792, April 1996. ISSN 2155-3165. doi : 10.1364/AO.35.001780.

[66] Markus G. Müller, Irene Georgakoudi, Qingguo Zhang, Jun Wu, and Michael S. Feld. Intrinsic fluorescence spectroscopy in turbid media : Disentangling effects of scattering and absorption. *Applied Optics*, 40(25) :4633–4646, September 2001. ISSN 2155-3165. doi : 10.1364/AO.40.004633.

[67] Sidan Tian, Zeyu Zhang, Fanling Meng, Zhihua Wang, and Liang Luo. Recent Advances in Enhancement of Raman Scattering Intensity for Biological Applications. *Chemical & Biomedical Imaging*, 1(7) :575–589, October 2023. doi : 10.1021/cbmi.3c00017.

[68] L.-P. Choo-Smith, H. G. M. Edwards, H. P. Endtz, J. M. Kros, F. Heule, H. Barr, J. S. Robinson Jr., H. A. Bruining, and G. J. Puppels. Medical applications of Raman spectroscopy : From proof of principle to clinical implementation. *Biopolymers*, 67(1) : 1–9, 2002. ISSN 1097-0282. doi : 10.1002/bip.10064.

[69] Eliana Cordero, Ines Latka, Christian Matthäus, Iwan W. Schie, and Jürgen Popp. [i]In-vivo[/i] Raman spectroscopy : From basics to applications. *Journal of Biomedical Optics*, 23(7) :071210, June 2018. ISSN 1083-3668, 1560-2281. doi : 10.1117/1.JBO.23.7.071210.

[70] Maria Anthi Kouri, Ellas Spyratou, Maria Karnachoriti, Dimitris Kalatzis, Nikolaos Danias, Nikolaos Arkadopoulos, Ioannis Seimenis, Yannis S. Raptis, Athanassios G. Kontos, and Efstathios P. Efstathopoulos. Raman Spectroscopy : A Personalized Decision-Making Tool on Clinicians' Hands for In Situ Cancer Diagnosis and Surgery Guidance. *Cancers*, 14(5) :1144, January 2022. ISSN 2072-6694. doi : 10.3390/cancers14051144.

[71] Zanyar Movasaghi, Rehman , Shazza, and Ihtesham U. and Rehman. Raman Spectroscopy of Biological Tissues. *Applied Spectroscopy Reviews*, 42(5) :493–541, September 2007. ISSN 0570-4928. doi : 10.1080/05704920701551530.

[72] Nikki Kuhar, Sanchita Sil, Taru Verma, and Siva Umapathy. Challenges in application of Raman spectroscopy to biology and materials. *RSC Advances*, 8(46) :25888–25908, July 2018. ISSN 2046-2069. doi : 10.1039/C8RA04491K.

[73] Luís F. Santos, Rolf Wolthuis, S. Koljenović, Rui M. Almeida, and Gerwin J. Puppels. Fiber-optic probes for in vivo Raman spectroscopy in the high-wavenumber region. *Analytical Chemistry*, 77(20) :6747–6752, October 2005. ISSN 0003-2700. doi : 10.1021/ac0505730.

[74] S. Koljenović, T. C. Bakker Schut, R. Wolthuis, A. J. P. E. Vincent, G. Hendriks-Hagevi, L. Santos, J. M. Kros, and G. J. Puppels. Raman spectroscopic characterization of porcine brain tissue using a single fiber-optic probe. *Analytical Chemistry*, 79(2) : 557–564, January 2007. ISSN 0003-2700. doi : 10.1021/ac0616512.

[75] Marie-Louise O'Connell, Alan G. Ryder, Marc N. Leger, and Tom Howley. Qualitative analysis using Raman spectroscopy and chemometrics : A comprehensive model system for narcotics analysis. *Applied Spectroscopy*, 64(10) :1109–1121, October 2010. ISSN 1943-3530. doi : 10.1366/000370210792973541.

[76] Qiang Tu and Chang Chang. Diagnostic applications of Raman spectroscopy. *Nanomedicine : Nanotechnology, Biology and Medicine*, 8(5) :545–558, July 2012. ISSN 1549-9634. doi : 10.1016/j.nano.2011.09.013.

[77] Chi Gao, Peng Zhao, Qi Fan, Haonan Jing, Ruochen Dang, Weifeng Sun, Yutao Feng, Bingliang Hu, and Quan Wang. Deep Neural Network : As the Novel Pipelines in Multiple Preprocessing for Raman Spectroscopy, March 2023.

[78] Andrew P. Shreve, Nerine J. Cherepy, and Richard A. Mathies. Effective Rejection of Fluorescence Interference in Raman Spectroscopy Using a Shifted Excitation Difference Technique. *Applied Spectroscopy*, 46(4) :707–711, April 1992. ISSN 0003-7028. doi : 10.1366/0003702924125122.

[79] Thomas Bocklitz, Angela Walter, Katharina Hartmann, Petra Rösch, and Jürgen Popp. How to pre-process Raman spectra for reliable and stable models ? *Analytica Chimica Acta*, 704(1-2) :47–56, October 2011. ISSN 1873-4324. doi : 10.1016/j.aca.2011.06.043.

[80] Chad A. Lieber and Anita Mahadevan-Jansen. Automated Method for Subtraction of Fluorescence from Biological Raman Spectra. *Applied Spectroscopy*, 57(11) :1363–1367, November 2003. ISSN 0003-7028. doi : 10.1366/000370203322554518.

[81] Jianhua Zhao, Harvey Lui, David I. McLean, and Haishan Zeng. Automated Autofluorescence Background Subtraction Algorithm for Biomedical Raman Spectroscopy. *Applied Spectroscopy*, 61(11) :1225–1232, November 2007. ISSN 0003-7028, 1943-3530. doi : 10.1366/000370207782597003.

[82] P. J. Cadusch, M. M. Hlaing, S. A. Wade, S. L. McArthur, and P. R. Stoddart. Improved methods for fluorescence background subtraction from Raman spectra. *Journal of Raman Spectroscopy*, 44(11) :1587–1595, 2013. ISSN 1097-4555. doi : 10.1002/jrs.4371.

[83] Brooke D. Beier and Andrew J. Berger. Method for automated background subtraction from Raman spectra containing known contaminants. *The Analyst*, 134(6) :1198–1202, June 2009. ISSN 1364-5528. doi : 10.1039/b821856k.

[84] Rosanna Perez-Pueyo, Maria Jose Soneira, and Sergio Ruiz-Moreno. Morphology-Based Automated Baseline Removal for Raman Spectra of Artistic Pigments. *Applied Spectroscopy*, 64(6) :595–600, June 2010. ISSN 0003-7028. doi : 10.1366/000370210791414281.

[85] Fabiola León-Bejarano, Miguel Ramírez-Elías, Martin O. Mendez, Guadalupe Dorantes-Méndez, Ma. del Carmen Rodríguez-Aranda, and Alfonso Alba. Denoising of Raman spectroscopy for biological samples based on empirical mode decomposition. *International Journal of Modern Physics C*, 28(09) :1750116, September 2017. ISSN 0129-1831. doi : 10.1142/S0129183117501169.

[86] Christoph Krafft, Benjamin Dietzek, and Jürgen Popp. Raman and CARSmicrospectroscopy of cells and tissues. *Analyst*, 134(6) :1046–1057, June 2009. ISSN 1364-5528. doi : 10.1039/B822354H.

[87] W. J. Niels Klement, Philippe Leproux, Wesley R. Browne, and Hideaki Kano. CMOS and CCD Detection in Raman Spectroscopy : A Comparison Using Spontaneous and Multiplex Coherent Anti-Stokes Raman Scattering (CARS). *Journal of Raman Spectroscopy*, n/a(n/a), 2025. ISSN 1097-4555. doi : 10.1002/jrs.6773.

[88] Eliana Cordero, Florian Korinth, Clara Stiebing, Christoph Krafft, Iwan W. Schie, and Jürgen Popp. Evaluation of Shifted Excitation Raman Difference Spectroscopy and Comparison to Computational Background Correction Methods Applied to Biochemical Raman Spectra. *Sensors*, 17(8) :1724, August 2017. ISSN 1424-8220. doi : 10.3390/s17081724.

[89] Marco Pinto Corujo, Pavel Michal, Rod Wesson, Don Praveen Amarasinghe, Alison Rodger, and Nikola P. Chmel. Reduction of Background Fluorescence from Impurities

in Protein Samples for Raman Spectroscopy. *Journal of Spectroscopy*, 2022(1) :1928091, 2022. ISSN 2314-4939. doi : 10.1155/2022/1928091.

[90] Tatu Rojalin, Lauri Kurki, Timo Laaksonen, Tapani Viitala, Juha Kostamovaara, Keith C. Gordon, Leonardo Galvis, Sebastian Wachsmann-Hogiu, Clare J. Strachan, and Marjo Yliperttula. Fluorescence-suppressed time-resolved Raman spectroscopy of pharmaceuticals using complementary metal-oxide semiconductor (CMOS) single-photon avalanche diode (SPAD) detector. *Analytical and Bioanalytical Chemistry*, 408 (3) :761–774, January 2016. ISSN 1618-2650. doi : 10.1007/s00216-015-9156-6.

[91] Evtim V. Efremov, Joost B. Buijs, Cees Gooijer, and Freek Ariese. Fluorescence Rejection in Resonance Raman Spectroscopy Using a Picosecond-Gated Intensified Charge-Coupled Device Camera. *Applied Spectroscopy*, 61(6) :571–578, June 2007. ISSN 0003-7028. doi : 10.1366/000370207781269873.

[92] Dong Wei, Chen , Shuo, and Quan and Liu. Review of Fluorescence Suppression Techniques in Raman Spectroscopy. *Applied Spectroscopy Reviews*, 50(5) :387–406, May 2015. ISSN 0570-4928. doi : 10.1080/05704928.2014.999936.

[93] P. Matousek, M. Towrie, and A. W. Parker. Fluorescence background suppression in Raman spectroscopy using combined Kerr gated and shifted excitation Raman difference techniques. *Journal of Raman Spectroscopy*, 33(4) :238–242, 2002. ISSN 1097-4555. doi : 10.1002/jrs.840.

[94] I. G. Cormack, M. Mazilu, K. Dholakia, and C. S. Herrington. Fluorescence suppression within Raman spectroscopy using annular beam excitation. *Applied Physics Letters*, 91(2) :023903, July 2007. ISSN 0003-6951. doi : 10.1063/1.2756311.

[95] Katharina Hering, Dana Cialla, Katrin Ackermann, Thomas Dörfer, Robert Möller, Henrik Schneidewind, Roland Mattheis, Wolfgang Fritzsche, Petra Rösch, and Jürgen Popp. SERS : A versatile tool in chemical and biochemical diagnostics. *Analytical and Bioanalytical Chemistry*, 390(1) :113–124, January 2008. ISSN 1618-2642, 1618-2650. doi : 10.1007/s00216-007-1667-3.

[96] D. Cialla-May, X.-S. Zheng, K. Weber, and J. Popp. Recent progress in surface-enhanced Raman spectroscopy for biological and biomedical applications : From cells to clinics. *Chemical Society Reviews*, 46(13) :3945–3961, 2017. ISSN 0306-0012, 1460-4744. doi : 10.1039/C7CS00172J.

[97] Stefan Harmsen, Matthew A Wall, Ruimin Huang, and Moritz F Kircher. Cancer imaging using surface-enhanced resonance Raman scattering nanoparticles. *Nature Protocols*, 12(7) :1400–1414, July 2017. ISSN 1754-2189, 1750-2799. doi : 10.1038/nprot.2017.031.

[98] Peter S. Millard, Juan Bru, and Christopher A. Berger. Open-source point-of-care electronic medical records for use in resource-limited settings : Systematic review and questionnaire surveys. *BMJ Open*, 2(4) :e000690, January 2012. ISSN 2044-6055, 2044-6055. doi : 10.1136/bmjopen-2011-000690.

[99] M. L. M. Kiah, Ahmed Haiqi, B. B. Zaidan, and A. A. Zaidan. Open source EMR software : Profiling, insights and hands-on analysis. *Computer Methods and Programs in Biomedicine*, 117(2) :360–382, November 2014. ISSN 0169-2607. doi : 10.1016/j.cmpb.2014.07.002.

[100] Osman Ratib and Antoine Rosset. Open-source software in medical imaging : Development of OsiriX. *International Journal of Computer Assisted Radiology and Surgery*, 1(4) :187–196, December 2006. ISSN 1861-6429. doi : 10.1007/s11548-006-0056-2.

[101] Osman Ratib, Antoine Rosset, and Joris Heuberger. Open Source software and social networks : Disruptive alternatives for medical imaging. *European Journal of Radiology*, 78(2) :259–265, May 2011. ISSN 0720-048X, 1872-7727. doi : 10.1016/j.ejrad.2010.05.004.

[102] Luciano Prevedello and Ramin Khorasani. Should You Use Open-Source Software Applications in Your Practice ? *Journal of the American College of Radiology*, 9(10) : 751–752, October 2012. ISSN 1546-1440, 1558-349X. doi : 10.1016/j.jacr.2012.06.033.

[103] Zhengru Shen and Marco Spruit. A Systematic Review of Open Source Clinical Software on GitHub for Improving Software Reuse in Smart Healthcare. *Applied Sciences*, 9(1) : 150, January 2019. ISSN 2076-3417. doi : 10.3390/app9010150.

[104] Deniz İlhan Topcu. Developing Data-Centric Clinical Laboratory Workflow Through the Use of Open-Source Tools. *The Journal of Applied Laboratory Medicine*, 8(1) :7–10, January 2023. ISSN 2576-9456. doi : 10.1093/jalm/jfac110.

[105] Nicole M. Ralbovsky and Igor K. Lednev. Towards development of a novel universal medical diagnostic method : Raman spectroscopy and machine learning. *Chemical Society Reviews*, 49(20) :7428–7453, October 2020. ISSN 1460-4744. doi : 10.1039/D0CS01019G.

[106] Carl Rafferty, Kjell Johnson, Jim O'Mahony, Barbara Burgoyne, Rosemary Rea, and Karin M. Balss. Analysis of chemometric models applied to Raman spectroscopy for monitoring key metabolites of cell culture. *Biotechnology Progress*, 36(4) :e2977, July 2020. ISSN 1520-6033. doi : 10.1002/btpr.2977.

[107] Nathan Blake, Riana Gaifulina, Lewis D. Griffin, Ian M. Bell, and Geraint M. H. Thomas. Machine Learning of Raman Spectroscopy Data for Classifying Cancers : A Review of the Recent Literature. *Diagnostics*, 12(6) :1491, June 2022. ISSN 2075-4418. doi : 10.3390/diagnostics12061491.

[108] S. Gomes da Costa, A. Richter, U. Schmidt, S. Breuninger, and O. Hollricher. Confocal Raman microscopy in life sciences. *Morphologie*, 103(341) :11–16, March 2019. ISSN 12860115. doi : 10.1016/j.morpho.2018.12.003.

[109] Razvigor Darlenski and Joachim W. Fluhr. In vivo Raman Confocal Spectroscopy in the Investigation of the Skin Barrier. In T. Agner, editor, *Current Problems in Dermatology*, volume 49, pages 71–79. S. Karger AG, 2016. ISBN 978-3-318-05585-6 978-3-318-05586-3. doi : 10.1159/000441587.

[110] Matthew J. Baker, Shawn R. Hussain, Lila Lovergne, Valérie Untereiner, Caryn Hughes, Roman A. Lukaszewski, Gérard Thiéfin, and Ganesh D. Sockalingum. Developing and understanding biofluid vibrational spectroscopy : A critical review. *Chemical Society Reviews*, 45(7) :1803–1818, 2016. ISSN 0306-0012, 1460-4744. doi : 10.1039/C5CS00585J.

[111] Michelle Brusatori, Gregory Auner, Thomas Noh, Lisa Scarpace, Brandy Broadbent, and Steven N. Kalkanis. Intraoperative Raman Spectroscopy. *Neurosurgery Clinics of North America*, 28(4) :633–652, October 2017. ISSN 10423680. doi : 10.1016/j.nec.2017.05.014.

[112] Todd Hollon, Spencer Lewis, Christian W. Freudiger, X. Sunney Xie, and Daniel A. Orringer. Improving the accuracy of brain tumor surgery via Raman-based technology. *Neurosurgical Focus*, 40(3) :E9, March 2016. ISSN 1092-0684. doi : 10.3171/2015.12.FOCUS15557.

[113] Karl St-Arnaud, Kelly Aubertin, Mathias Strupler, Wendy-Julie Madore, Andrée-Anne Grosset, Kevin Petrecca, Dominique Trudel, and Frédéric Leblond. Development and characterization of a handheld hyperspectral Raman imaging probe system for molecular characterization of tissue on mesoscopic scales. *Medical Physics*, 45(1) :328–339, January 2018. ISSN 00942405. doi : 10.1002/mp.12657.

[114] François Daoust, Tien Nguyen, Patrick Orsini, Jacques Bismuth, Marie-Maude de Denus-Baillargeon, Israel Veilleux, Alexandre Wetter, Philippe McKoy, Isabelle Dicaire, Maroun Massabki, Kevin Petrecca, and Frédéric Leblond. Handheld macroscopic Raman spectroscopy imaging instrument for machine-learning-based molecular tissue margins characterization. *Journal of Biomedical Optics*, 26(2) :022911, February 2021. ISSN 1083-3668, 1560-2281. doi : 10.1117/1.JBO.26.2.022911.

[115] Ruochong Zhang, Renzhe Bi, Chris Ho Jun Hui, Poongkulali Rajarahm, U. S. Dinish, and Malini Olivo. A Portable Ultrawideband Confocal Raman Spectroscopy System with a Handheld Probe for Skin Studies. *ACS sensors*, 6(8) :2960–2966, August 2021. ISSN 2379-3694. doi : 10.1021/acssensors.1c00761.

[116] Adrian Ghita, Pavel Matousek, and Nick Stone. Sensitivity of Transmission Raman Spectroscopy Signals to Temperature of Biological Tissues. *Scientific Reports*, 8(1) : 8379, December 2018. ISSN 2045-2322. doi : 10.1038/s41598-018-25465-x.

[117] Sara Mosca, Priyanka Dey, Marzieh Salimi, Benjamin Gardner, Francesca Palombo, Nick Stone, and Pavel Matousek. Spatially Offset Raman Spectroscopy—How Deep ? *Analytical Chemistry*, 93(17) :6755–6762, May 2021. ISSN 0003-2700. doi : 10.1021/acs.analchem.1c00490.

[118] Pavel Matousek and Nicholas Stone. Development of deep subsurface Raman spectroscopy for medical diagnosis and disease monitoring. *Chemical Society Reviews*, 45(7) : 1794–1802, 2016. ISSN 0306-0012, 1460-4744. doi : 10.1039/C5CS00466G.

[119] P. Matousek, I. P. Clark, E. R. C. Draper, M. D. Morris, A. E. Goodship, N. Everall, M. Towrie, W. F. Finney, and A. W. Parker. Subsurface Probing in Diffusely Scattering Media Using Spatially Offset Raman Spectroscopy. *Applied Spectroscopy*, 59(4) :393–400, April 2005. ISSN 0003-7028, 1943-3530. doi : 10.1366/0003702053641450.

[120] P. Matousek, M. D. Morris, N. Everall, I. P. Clark, M. Towrie, E. Draper, A. Goodship, and A. W. Parker. Numerical Simulations of Subsurface Probing in Diffusely Scattering Media Using Spatially Offset Raman Spectroscopy. *Applied Spectroscopy*, 59(12) :1485–1492, December 2005. ISSN 0003-7028, 1943-3530. doi : 10.1366/000370205775142548.

[121] Chi Shu, Keren Chen, Maria Lynch, Jason R. Maher, Hani A. Awad, and Andrew J. Berger. Spatially offset Raman spectroscopy for in vivo bone strength prediction. *Biomedical Optics Express*, 9(10) :4781–4791, October 2018. ISSN 2156-7085. doi : 10.1364/BOE.9.004781.

[122] Fay Nicolson, Moritz F. Kircher, Nick Stone, and Pavel Matousek. Spatially offset Raman spectroscopy for biomedical applications. *Chemical Society Reviews*, 50(1) : 556–568, 2021. ISSN 0306-0012, 1460-4744. doi : 10.1039/D0CS00855A.

[123] Judith Langer, Dorleta Jimenez de Aberasturi, Javier Aizpurua, Ramon A. Alvarez-Puebla, Baptiste Auguié, Jeremy J. Baumberg, Guillermo C. Bazan, Steven E. J. Bell, Anja Boisen, Alexandre G. Brolo, Jaebum Choo, Dana Cialla-May, Volker Deckert, Laura Fabris, Karen Faulds, F. Javier García de Abajo, Royston Goodacre, Duncan Graham, Amanda J. Haes, Christy L. Haynes, Christian Huck, Tamitake Itoh, Mikael Käll, Janina Kneipp, Nicholas A. Kotov, Hua Kuang, Eric C. Le Ru, Hiang Kwee Lee, Jian-Feng Li, Xing Yi Ling, Stefan A. Maier, Thomas Mayerhöfer, Martin Moskovits, Kei Murakoshi, Jwa-Min Nam, Shuming Nie, Yukihiro Ozaki, Isabel Pastoriza-Santos, Jorge Perez-Juste, Juergen Popp, Annemarie Pucci, Stephanie Reich, Bin Ren, George C. Schatz, Timur Shegai, Sebastian Schlücker, Li-Lin Tay, K. George Thomas, Zhong-Qun Tian, Richard P. Van Duyne, Tuan Vo-Dinh, Yue Wang, Kathe-

rine A. Willets, Chuanlai Xu, Hongxing Xu, Yikai Xu, Yuko S. Yamamoto, Bing Zhao, and Luis M. Liz-Marzán. Present and Future of Surface-Enhanced Raman Scattering. *ACS Nano*, 14(1) :28–117, January 2020. ISSN 1936-0851, 1936-086X. doi : 10.1021/acsnano.9b04224.

[124] Pilot, Signorini, Durante, Orian, Bhamidipati, and Fabris. A Review on Surface-Enhanced Raman Scattering. *Biosensors*, 9(2) :57, April 2019. ISSN 2079-6374. doi : 10.3390/bios9020057.

[125] Yunqing Wang, Bing Yan, and Lingxin Chen. SERS Tags : Novel Optical Nanoprobes for Bioanalysis. *Chemical Reviews*, 113(3) :1391–1428, March 2013. ISSN 0009-2665, 1520-6890. doi : 10.1021/cr300120g.

[126] Medhanie Tesfay Gebrekidan, Ramona Erber, Arndt Hartmann, Peter A. Fasching, Julius Emons, Mathias W. Beckmann, and Andreas Braeuer. Breast Tumor Analysis Using Shifted-Excitation Raman Difference Spectroscopy (SERDS). *Technology in Cancer Research & Treatment*, 17 :153303381878253, January 2018. ISSN 1533-0346, 1533-0338. doi : 10.1177/1533033818782532.

[127] F. Korinth, A. S. Mondol, C. Stiebing, I. W. Schie, C. Krafft, and J. Popp. New methodology to process shifted excitation Raman difference spectroscopy data : A case study of pollen classification. *Scientific Reports*, 10(1) :11215, December 2020. ISSN 2045-2322. doi : 10.1038/s41598-020-67897-4.

[128] Florian Korinth, Tanveer Ahmed Shaik, Jürgen Popp, and Christoph Krafft. Assessment of shifted excitation Raman difference spectroscopy in highly fluorescent biological samples. *The Analyst*, 146(22) :6760–6767, November 2021. ISSN 1364-5528. doi : 10.1039/d1an01376a.

[129] Steven J. Choquette, Edgar S. Etz, Wilbur S. Hurst, Douglas H. Blackburn, and Stefan D. Leigh. Relative intensity correction of Raman spectrometers : NIST SRMs 2241 through 2243 for 785 nm, 532 nm, and 488 nm/514.5 nm excitation. *Applied Spectroscopy*, 61(2) :117–129, February 2007. ISSN 0003-7028. doi : 10.1366/000370207779947585.

[130] Nassim Ksantini, Israel Veilleux, Marie-Maude de Denus-Baillargeon, Patrick Orsini, Isabelle Dicaire, Jean-Bernard Lecourt, Alexandre Gognau, Yves Hernandez, Antonio Baylon, Maroun Massabki, Frederic Lesage, and Frederic Leblond. Time-gated interferometric detection increases Raman scattering to fluorescence signal ratio in biological samples. *Journal of Biophotonics*, 15(1) :e202100188, January 2022. ISSN 1864-0648. doi : 10.1002/jbio.202100188.

[131] Boyan Li, Amandine Calvet, Yannick Casamayou-Boucau, and Alan G. Ryder. Kernel principal component analysis residual diagnosis (KPCARD) : An automated method

for cosmic ray artifact removal in Raman spectra. *Analytica Chimica Acta*, 913 :111–120, March 2016. ISSN 00032670. doi : 10.1016/j.aca.2016.01.042.

[132] Stephen M. Anthony and Jerilyn A. Timlin. Removing Cosmic Spikes Using a Hyperspectral Upper-Bound Spectrum Method. *Applied Spectroscopy*, 71(3) :507–519, March 2017. ISSN 0003-7028, 1943-3530. doi : 10.1177/0003702816668528.

[133] Sinead J. Barton and Bryan M. Hennelly. An Algorithm for the Removal of Cosmic Ray Artifacts in Spectral Data Sets. *Applied Spectroscopy*, 73(8) :893–901, August 2019. ISSN 0003-7028, 1943-3530. doi : 10.1177/0003702819839098.

[134] Holly J. Butler, Lorna Ashton, Benjamin Bird, Gianfelice Cinque, Kelly Curtis, Jennifer Dorney, Karen Esmonde-White, Nigel J. Fullwood, Benjamin Gardner, Pierre L. Martin-Hirsch, Michael J. Walsh, Martin R. McAinsh, Nicholas Stone, and Francis L. Martin. Using Raman spectroscopy to characterize biological materials. *Nature Protocols*, 11(4) :664–687, April 2016. ISSN 1750-2799. doi : 10.1038/nprot.2016.036.

[135] Fabien Picot, François Daoust, Guillaume Sheehy, Frédérick Dallaire, Layal Chaikho, Théophile Bégin, Samuel Kadoury, and Frédéric Leblond. Data consistency and classification model transferability across biomedical Raman spectroscopy systems. *Translational Biophotonics*, 3(1), March 2021. ISSN 2627-1850, 2627-1850. doi : 10.1002/tbio.202000019.

[136] Isaac O. Afara, Rubina Shaikh, Ervin Nippolainen, William Querido, Jari Torniainen, Jaakko K. Sarin, Shital Kandel, Nancy Pleshko, and Juha Töyräs. Characterization of connective tissues using near-infrared spectroscopy and imaging. *Nature Protocols*, 16(2) :1297–1329, February 2021. ISSN 1754-2189, 1750-2799. doi : 10.1038/s41596-020-00468-z.

[137] Frédérick Dallaire, Fabien Picot, Jean-Philippe Tremblay, Guillaume Sheehy, Émile Lemoine, Rajeev Agarwal, Samuel Kadoury, Dominique Trudel, Frédéric Lesage, Kevin Petrecca, and Frédéric Leblond. Quantitative spectral quality assessment technique validated using intraoperative in vivo Raman spectroscopy measurements. *Journal of Biomedical Optics*, 25(04) :1, April 2020. ISSN 1083-3668. doi : 10.1117/1.JBO.25.4.040501.

[138] ANSI. *American National Standard for Safe Use of Lasers*. Number Z136.1-2014. ANSI, Orlando, FL, 2014. ISBN 978-1-940168-01-2.

[139] ANSI. *American National Standard for Safe Use of Lasers in Health Care*. Number Z136.3-2011. ANSI, Orlando, FL, 2011. ISBN 978-0-912035-69-7.

[140] Eric T. Marple. Filtered fiber optic probe, May 2012.

[141] Eric T. Marple and Kirk D. Urmey. Fiber optic probes utilizing GRIN lenses for spatially precise optical spectroscopy, April 2013.

[142] Eric T. Marple and Kirk D. Urmey. Filtered fiber optic probe, April 2014.

[143] Eric Todd Marple and Kirk David Urmey. Multi-fiber optical connector assembly that includes an optical connector and a connector housing, August 2017.

[144] Eric Todd Marple, Kirk David Urmey, and John Meckert. Spectrometer having a variable focus lens with a single air gap, October 2019.

[145] Eric Todd Marple, Kirk David Urmey, and John Meckert. Spectrometers having a variable focus lens with a single air gap, November 2020.

[146] Guillaume Sheehy. Development of a Spatial Offset Raman Spectroscopy (SORS) Platform for Biological Tissue Interrogation. Master's thesis, École Polytechnique de Montréal, August 2018.

[147] Guillaume Sheehy, Fabien Picot, Frédérick Dallaire, Katherine Ember, Tien Nguyen, Kevin Petrecca, Dominique Trudel, and Frédéric Leblond. Open-sourced Raman spectroscopy data processing package implementing a baseline removal algorithm validated from multiple datasets acquired in human tissue and biofluids. *Journal of Biomedical Optics*, 28(2) :025002, February 2023. ISSN 1560-2281. doi : 10.1117/1.JBO.28.2.025002.

[148] P McAllister and J Jeswiet. Medical device regulation for manufacturers. *Proceedings of the Institution of Mechanical Engineers, Part H : Journal of Engineering in Medicine*, 217(6) :459–467, June 2003. ISSN 0954-4119, 2041-3033. doi : 10.1243/09544110360729090.

[149] Yashika Sharma, Dilpreet Singh, and Amrinder Singh. An Introduction to Regulatory Aspects Related to Medical Devices in Japan andCanada. *Current Indian Science*, 01, August 2023. ISSN 2210-299X. doi : 10.2174/2210299x01666230720103101.

[150] *Health Technology Assessment of Medical Devices*, pages 795–798. Elsevier, 2020. ISBN 978-0-12-813467-2. doi : 10.1016/b978-0-12-813467-2.00112-7.

[151] OSI. The Open Source AI Definition – 1.0, October 2024.

[152] Frédéric Leblond, Frédérick Dallaire, Katherine Ember, Alice Le Moël, Victor Blanquez-Yeste, Hugo Tavera, Guillaume Sheehy, Trang Tran, Marie-Christine Guiot, Alexander G. Weil, Roy Dudley, Costas Hadjipanayis, and Kevin Petrecca. Quantitative assessment of the generalizability of a brain tumor Raman spectroscopy machine learning model to various tumor types including astrocytoma and oligodendroglioma. *Journal of Biomedical Optics*, 30(1) :010501, January 2025. ISSN 1083-3668, 1560-2281. doi : 10.1117/1.JBO.30.1.010501.

[153] Katherine J. I. Ember, Nassim Ksantini, Frédérick Dallaire, Guillaume Sheehy, Trang Tran, Mathieu Dehaes, Madeleine Durand, Dominique Trudel, and Frédéric Leblond.

Liquid saliva-based Raman spectroscopy device with on-board machine learning detects COVID-19 infection in real-time. *Analyst*, 149(22) :5535–5545, November 2024. ISSN 1364-5528. doi : 10.1039/D4AN00729H.

[154] Esmat Zamani, Nassim Ksantini, Guillaume Sheehy, Katherine J. I. Ember, Bill Baloukas, Oleg Zabeida, Tran Trang, Myriam Mahfoud, Jolanta-Ewa Sapieha, Ludvik Martinu, and Frédéric Leblond. Spectral effects and enhancement quantification in healthy human saliva with surface-enhanced Raman spectroscopy using silver nanopillar substrates. *Lasers in Surgery and Medicine*, 56(2) :206–217, February 2024. ISSN 1096-9101. doi : 10.1002/lsm.23746.

[155] Sandryne David, Arthur Plante, Frédérick Dallaire, Jean-Philippe Tremblay, Guillaume Sheehy, Elizabeth Macdonald, Laura Forrest, Manijeh Daneshmand, Dominique Trudel, Brian C. Wilson, Laura Hopkins, Sangeeta Murugkar, Barbara Vanderhyden, and Frédéric Leblond. Multispectral label-free Raman spectroscopy can detect ovarian and endometrial cancer with high accuracy. *Journal of Biophotonics*, 15(2) :e202100198, 2022. ISSN 1864-0648. doi : 10.1002/jbio.202100198.

[156] Roozbeh Shams, Fabien Picot, David Grajales, Guillaume Sheehy, Frederick Dallaire, Mirela Birlea, Fred Saad, Dominique Trudel, Cynthia Menard, Frederic Leblond, and Samuel Kadoury. Pre-clinical evaluation of an image-guided in-situ Raman spectroscopy navigation system for targeted prostate cancer interventions. *International Journal of Computer Assisted Radiology and Surgery*, 15(5) :867–876, May 2020. ISSN 1861-6429. doi : 10.1007/s11548-020-02136-9.

[157] Alireza Akbarzadeh, Ehsan Edjlali, Guillaume Sheehy, Juliette Selb, Rajeev Agarwal, Jessie Weber, and Frédéric Leblond. Experimental validation of a spectroscopic Monte Carlo light transport simulation technique and Raman scattering depth sensing analysis in biological tissue. *Journal of Biomedical Optics*, 25(10), October 2020. ISSN 1560-2281. doi : 10.1117/1.JBO.25.10.105002.

[158] Emile Beaulieu. Wide-Field Spatial Frequency Domain Imaging, Diffuse Reflectance and Endogenous Fluorescence Spectroscopy System for Quantitative Tissue Biomarkers in Radical Prostatectomy Specimens. Master's thesis, Polytechnique Montréal, August 2019.

[159] Hancong Liu, Sirish Shah, and Wei Jiang. On-line outlier detection and data cleaning. *Computers & Chemical Engineering*, 28(9) :1635–1647, August 2004. ISSN 0098-1354. doi : 10.1016/j.compchemeng.2004.01.009.

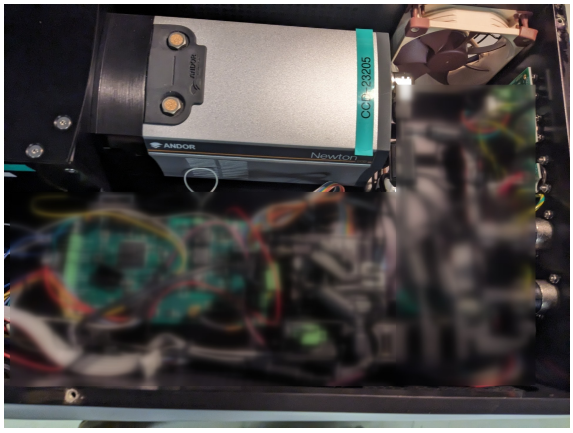
[160] Jukka Suomela. Median Filtering is Equivalent to Sorting, June 2014.

APPENDIX A ASSEMBLY GUIDE FOR THE LUMED RAMAN SYSTEM

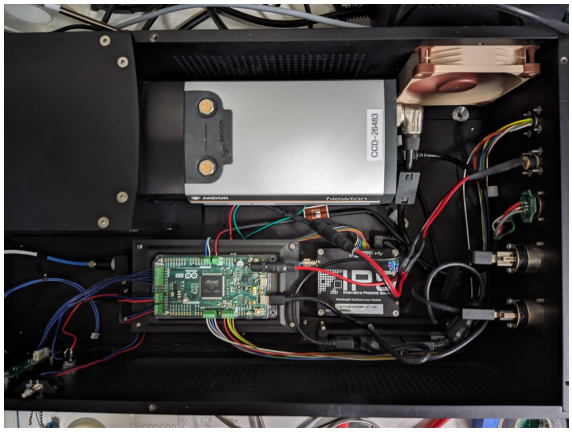
This appendix presents a detailed step-by-step guide for upgrading an ODS/Reveal Sentry unit to the Lumed Raman System. The process includes hardware modifications, power supply changes, electronic rewiring, and integration of the Lumed Raman hardware and control software stack.

Before and After Comparison

Figure A.1 illustrates the system configuration before and after the upgrade. The comparison highlights the extent of the hardware and wiring modifications required to transition from the original ODS/Reveal setup to the Lumed Raman System configuration.



(a) Original ODS/Reveal system configuration



(b) Upgraded Lumed Raman System configuration

FIGURE A.1 Comparison of the internal electronics and wiring configuration before and after the system upgrade

Removing Existing Components

Begin the upgrade process by removing all pre-existing hardware related to the ODS/Reveal configuration. This includes the laser control printed circuit board (PCB), the associated wiring harnesses, and the integrated USB hub if present.

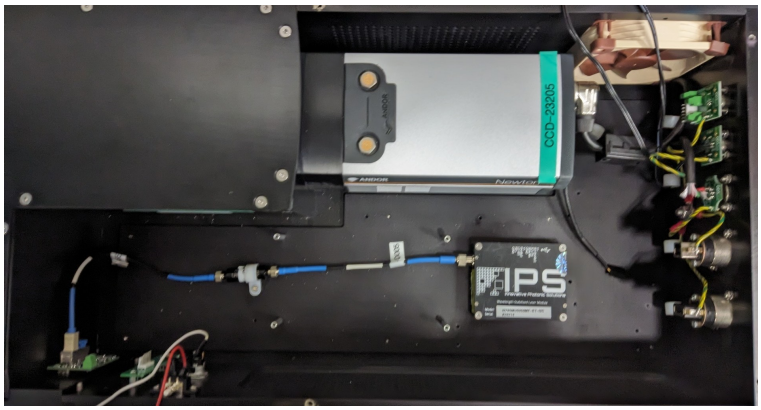


FIGURE A.2 Interior view of the system case after removal of original components, wiring, and USB hub.

Changing the Power Supply

The power supply connected to the PSU-2 port of the original ODS/Reveal Sentry unit must be replaced to accommodate the operational requirements of the Lumed Raman System. Specifically, the existing 18 V power supply must be replaced with a 12 V, 5 A DC power brick. There are two viable ways for this modification, depending on whether one intends to keep the original connector plug.

— **Option 1 :** Keep the existing screw-type 2-pin connector that is pre-installed in the ODS Sentry unit case, and modify the new 12 V PSU.

— **Option 2 :** Replace the existing connector on the Sentry case with a compatible panel-mount DC jack, allowing the new 12 V power adapter to be used without modification.

> **Tip : Option 1** requires fine motor skills and advanced soldering proficiency to ensure proper electrical contact and safety. Although **Option 2** is significantly easier to implement, it results in a slip-fit (non-screw) DC connector plug, which may be less mechanically secure.

Required Materials

— 1 × Power Supply, 12 V, 5 A DC (60 W) Adapter, 100 V to 240 V AC input

— 1 × 5.5 mm × 2.1 mm Female DC Power Jack, 3-pin Panel Mount (for **Option 2**)

— 1 × 5.5 mm × 2.1 mm Male DC Pigtail Connector (for **Option 2**)

— 2 × 5.5 mm × 2.1 mm Male DC Pigtail Connectors (for Arduino and IPS laser connections)

Option 1 – Installing the Screw-Type 2-Pin Connector on a 12 V Power Supply

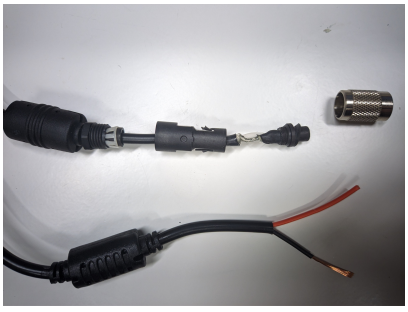
The Reveal/ODS Sentry unit employs a non-standard 2-pin M9-type screw connector for its PSU-2 port. This connector, manufactured by Binder [1], is designed for subminiature circular connections and offers a robust, secure locking mechanism. Unfortunately, it is not widely available in Canada and typically must be imported directly from the manufacturer.

To adapt a 12 V DC power supply for use with this connector, one of the following approaches may be used :

— Disassemble the original M9 connector from the 18 V power supply and reattach it to the new 12 V supply.
— Perform a cable splice by cutting and connecting the wires of the original PSU (with M9 connector) and the new 12 V PSU.

This process requires careful attention to polarity, insulation, and mechanical strain relief. It is recommended only for users with experience in cable termination and soldering.

(a) New 12 V PSU fitted with connector compatible with the Sentry PSU-2 port.



(b) Disassembled M9 connector from the 18 V PSU (above), and the cable from the 12 V PSU (below).

(c) M9 connector soldered to the 12 V PSU cable prior to final tightening and strain relief.

FIGURE A.3 Adaptation of a new 12 V power supply using a recycled M9 screw-type connector for compatibility with the original Sentry enclosure.

Option 2 – Installing a 5.5 mm × 2.1 mm Female Barrel Jack Connector

As an alternative to reusing the original M9 connector, the PSU-2 mounting plate can be replaced with a 5.5 mm × 2.1 mm female barrel jack connector. This approach eliminates

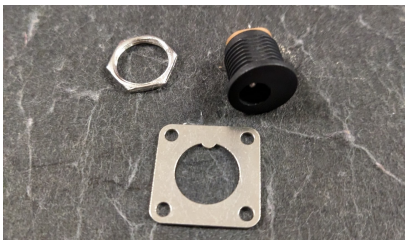
1. https://www.binder-connector.com/uk/products/subminiature-circular-connectors/m9-ip67/99-0402-00-02-m9-ip67-female-cable-connector-contacts-2-35-50-mm-unshielded-solder-ip67

the need for specialized soldering and enables compatibility with readily available 12 V DC power supplies.

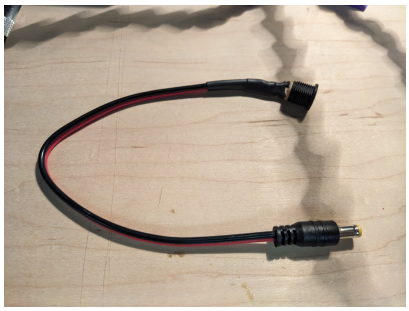
To implement this option :

1. Remove the stainless-steel mounting plate that houses the M9 connector.
2. Replace it with a new plate or modify the existing one to accommodate the panel-mount barrel jack connector.

(a) Original back panel showing the factory-installed M9 connector.



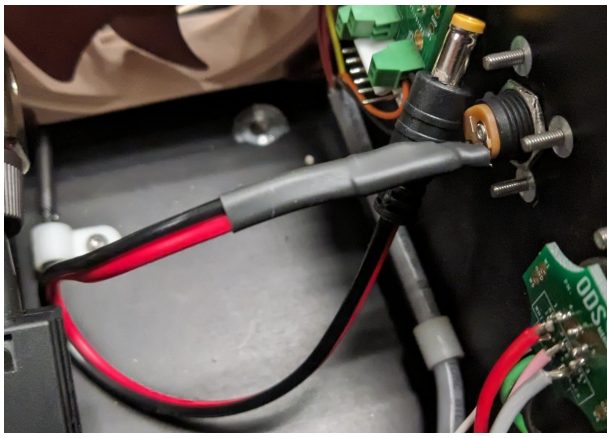
(b) Barrel jack connector (top) and PSU-2 mounting plate (bottom).



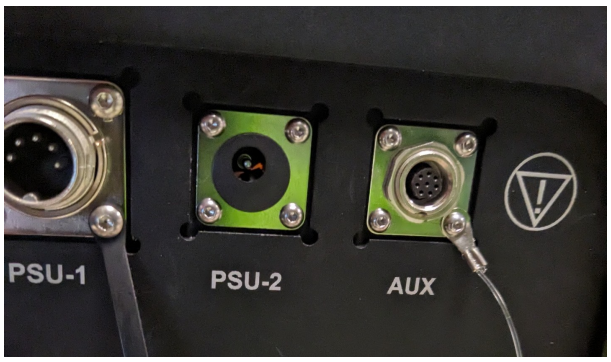
(c) Assembled barrel jack connector prior to installation.

FIGURE A.4 Components and preparation for installing a barrel jack PSU-2 connector in the Sentry unit.

> **Tip :** Soldering a male barrel pigtail to the panel-mount connector, as shown above, can simplify future servicing and reduce strain on the internal wiring. Alternatively, the power cable assembly may be soldered directly to the panel-mount connector (see Section A).



(a) Internal view of the unit following installation of the barrel jack connector.

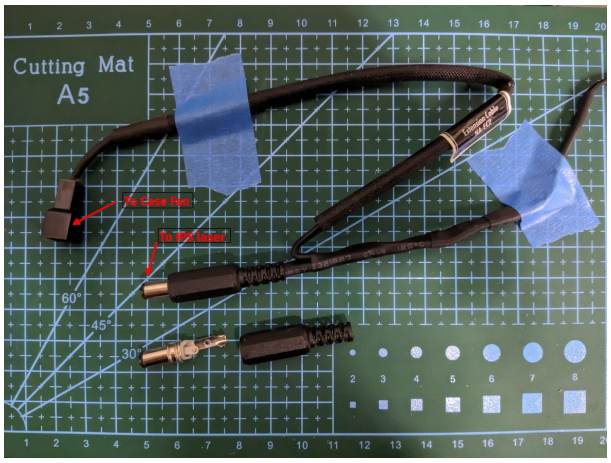


(b) External view of the new back panel with the installed barrel connector.

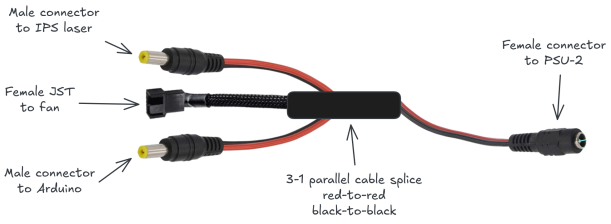
FIGURE A.5 Final installation views of the barrel jack PSU-2 connector.

Arduino, Laser, and Case Fan Connector

Power is distributed internally from the PSU-2 barrel jack connector to the Arduino microcontroller, the IPS laser module, and the internal cooling fan. These components share a common power line, assembled as a custom 3-to-1 power wiring assembly (Figure A.6b).



(a) **Previous version** of the internal power cable showing the fan wired in parallel with the laser connector.



(b) **Current version** of the power cable assembly.

FIGURE A.6 Evolution of the internal power distribution cable from PSU-2 to system components.

Laser Ribbon Connector

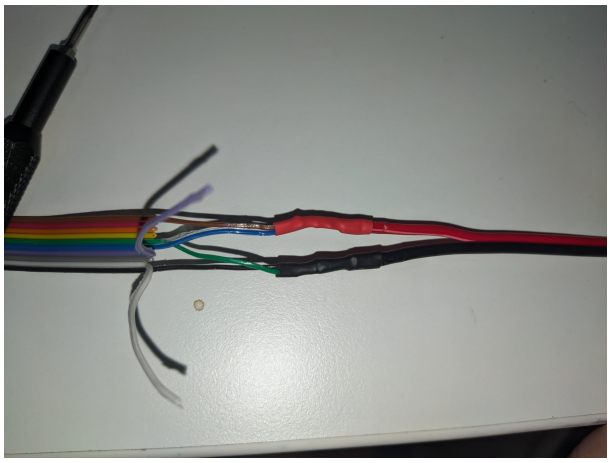
The IPS laser module receives power from the 12 V PSU-2 input and communicates via a Mini-USB interface. A custom ribbon cable is required to deliver power and signal connections between the laser head and the control electronics.

Required Materials

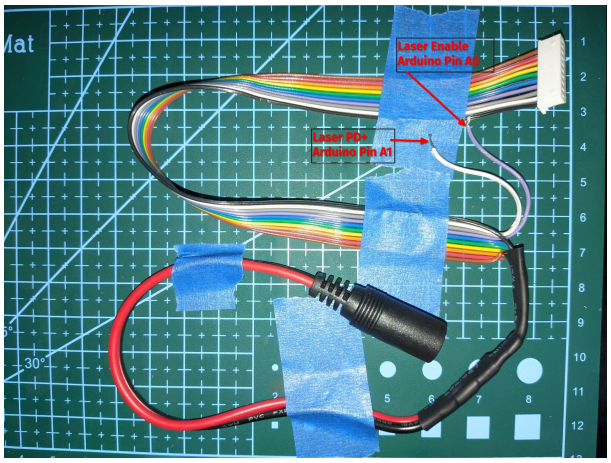
— 1 × IPS laser-compatible ribbon cable
— 1 × Female ribbon header connector (suitable for the IPS module pinout)

Wiring Configuration

When operated via Mini-USB, only the power pins on the ribbon cable require connection to the internal 12 V supply. Care must be taken to correctly identify the positive and ground lines according to the IPS manufacturer's specifications. Soldering should be followed by proper insulation, such as heat-shrink tubing, to ensure safety and reliability.



(a) Power pins of the IPS ribbon cable soldered and insulated with heat-shrink tubing.



(b) Completed IPS ribbon cable assembly.

FIGURE A.7 Preparation and assembly of the IPS laser ribbon power connector.

USB Hub Assembly

The USB hub consolidates USB connection from the different instruments (i.e. Laser, Camera and Powermeter) into a single cable connection between the Lumed Raman System and the

control computer. It is mounted internally using a custom 3D-printed fixture. A slim-form USB hub is recommended to ensure compatibility with the enclosure dimensions and ease of cable management.

Required Materials

— 1 × Anker Slim USB hub (or equivalent low-profile powered USB hub)

— Double-sided adhesive tape (e.g., 3M VHB) for securing the hub to the 3D-printed fixture

Installation Procedure

The USB hub is inserted into a 3D-printed mounting bracket specifically designed to fit the internal layout of the Lumed Raman System unit. The hub is fixed in place using double-sided adhesive tape to ensure mechanical stability in the mounting bracket.

(a) USB hub next to the 3D-printed bracket prior to assembly.

(b) USB hub secured in place within the 3D-printed fixture using adhesive tape.

FIGURE A.8 Assembly of the Anker slim USB hub and the 3D-printed internal fixture.

Arduino Assembly

The Lumed Raman System uses an Arduino Due microcontroller to coordinate signal routing, sensor interfacing, and hardware control. For mechanical and wiring flexibility, the version without pre-installed headers is mounted onto a custom 3D-printed fixture within the system enclosure. This configuration allows the use of terminal blocks for secure, serviceable

connections.

Required Materials

— $1\times$ Arduino Due (without headers)
— $1\times$ Right-angle through-hole DC power barrel jack
— $1\times$ 5.5 mm $\times$ 2.1 mm Female DC Power Jack, 3-pin Panel Mount
— $1\times$ 2-position terminal block
— $2\times$ 3-position terminal blocks
— $2\times$ 4-position terminal blocks

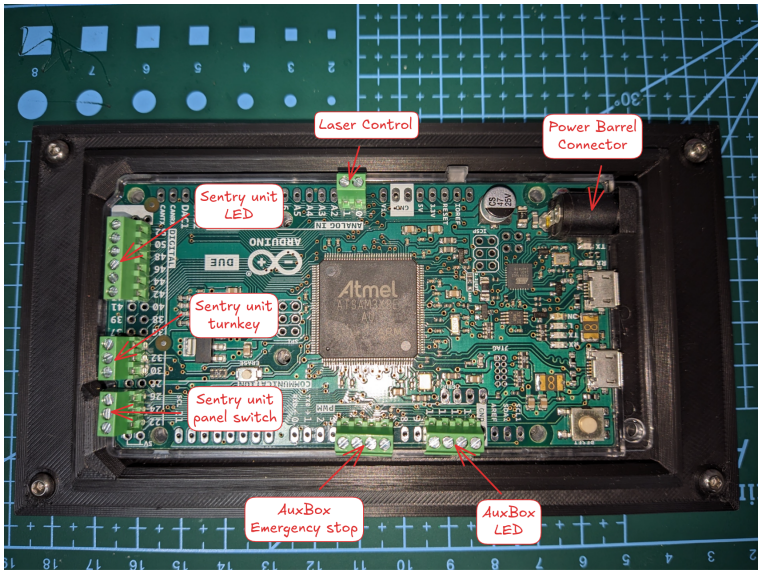
Pin Mapping and Terminal Block Connections

tableau A.1 summarizes the digital and analog pin usage of the Arduino Due, along with variable names used in the `.ino` source code, terminal block assignments, and functional descriptions of each connection.

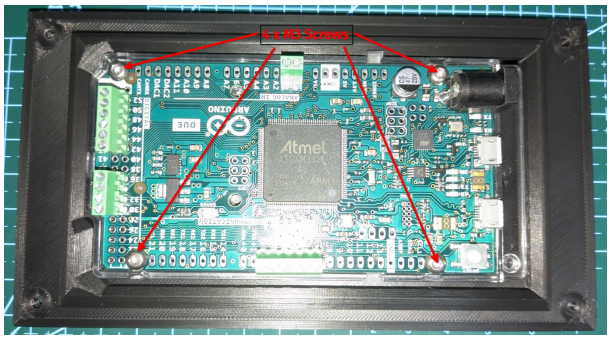
Pin	Variable (in .ino code)	Terminal Block	Description
A0	pin_ips_pdp	Laser Control	Analog input from IPS laser PD+ pin
A1	pin_ips_enable	Laser Control	Digital output to enable the laser (LOW or HIGH)
43	pin_unit_green_control	Sentry Unit LED Pins	GREEN LED control output
45	pin_unit_green_gnd	Sentry Unit LED Pins	GREEN LED ground connection (LOW)
47	pin_unit_yellow_control	Sentry Unit LED Pins	YELLOW LED control output
49	pin_unit_yellow_gnd	Sentry Unit LED Pins	YELLOW LED ground connection (LOW)
51	pin_unit_red_control	Sentry Unit LED Pins	RED LED control output
53	pin_unit_red_gnd	Sentry Unit LED Pins	RED LED ground connection (LOW)
35	pin_turnkey_ground	Sentry Unit Turnkey	Ground output for the turnkey switch (LOW)
33	pin_turnkey_probe	Sentry Unit Turnkey	Input probe for turnkey state (digitalRead)
31	pin_turnkey_high	Sentry Unit Turnkey	HIGH output for turnkey switch
27	pin_caseswitch_ground	Case Panel Switch	Ground output for panel switch (LOW)
25	pin_caseswitch_probe	Case Panel Switch	Input probe for panel switch (digitalRead)
23	pin_caseswitch_high	Case Panel Switch	HIGH output for panel switch
13	pin_aux_ground	AuxBox LED	Ground output for AuxBox LED (LOW)
12	pin_aux_green	AuxBox LED	GREEN AuxBox LED control output
11	pin_aux_yellow	AuxBox LED	YELLOW AuxBox LED control output
10	pin_aux_red	AuxBox LED	RED AuxBox LED control output
7	pin_aux_button_source	AuxBox Emergency Button	HIGH output for emergency button
6	pin_aux_button_probe	AuxBox Emergency Button	Input probe for emergency button (digitalRead)
5	pin_aux_grey	AuxBox Emergency Button	Not connected
4	pin_aux_white	AuxBox Emergency Button	Not connected

TABLE A.1 Arduino Due pin usage, variable assignments, terminal block layout, and functional roles.

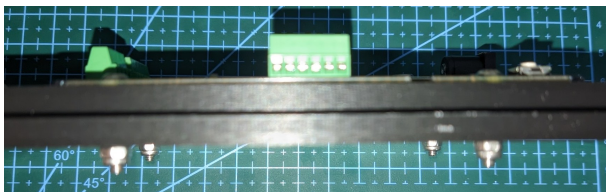
figure A.9 shows the layout and mechanical installation of the Arduino Due and associated terminal blocks on a 3D-printed fixture plate.



(a) Terminal block layout and pin distribution on the Arduino Due.



(b) Arduino Due mounted on a custom 3D-printed fixture plate.



(c) Side view showing embedded M3 nuts for mechanical mounting.

FIGURE A.9 Mechanical integration and terminal access for the Arduino Due assembly.

Switch Wiring

Switches in the Lumed Raman System are interfaced with the Arduino Due using standard digital input configurations. These include either 2-leg or 3-leg connection methods depending on the physical switch configuration and desired signal behavior.

Required Materials

— Solid-core 22 AWG wire
— 1 kΩ resistors
— Heat-shrink tubing

Option 1 : Two-Leg Switch Readout Circuit

In a two-leg switch configuration, the correct circuit involves placing a resistor and the switch in series between a constant positive voltage (e.g., 5 V) and ground. The Arduino input pin is connected at the junction between the switch and the resistor. This configuration ensures that :

— The input reads `HIGH` when the switch is open.
— The input reads `LOW` when the switch is closed.

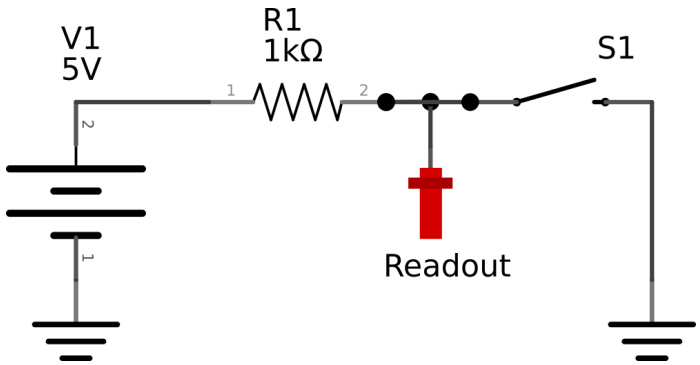
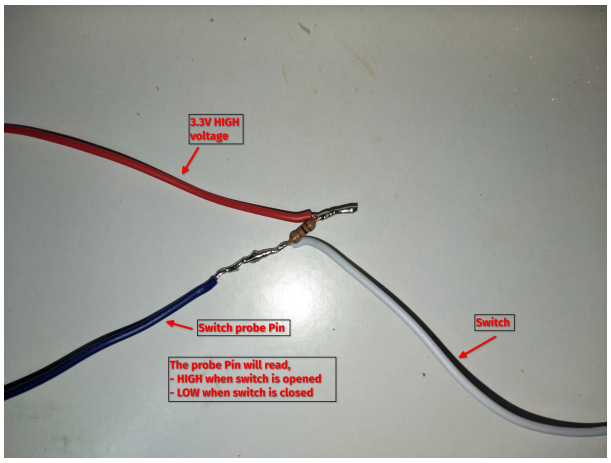


FIGURE A.10 Schematic for reading a 2-leg switch using a pull-down resistor and Arduino digital input.

Switch Resistor-Cable Assembly

To facilitate installation, the pull-down resistor may be integrated directly into the wiring harness. figure A.11 illustrates the soldered resistor concealed beneath transparent heat-shrink tubing for inspection and protection.



(a) Custom cable assembly showing soldered resistor embedded in wiring harness.

(b) Heat-shrink insulation over resistor and leads.

FIGURE A.11 Two-leg switch cable assembly with inline resistor.

In this configuration, the white wire connects to one leg of the switch, and the ground (e.g., black) wire to the other. The readout (blue) wire connects between the switch and the resistor, yielding a `HIGH` signal when the switch is open, and `LOW` when closed.

Option 2 : Three-Leg Switch Readout Circuit

For switches with three legs, a simplified wiring scheme may be used. The ground, probe (Arduino digital input), and 5 V high pins can be wired directly without an external pull-down resistor, as the Arduino pin current is limited. However, a 1 kΩ resistor may be added in series with the high-side line for protection.

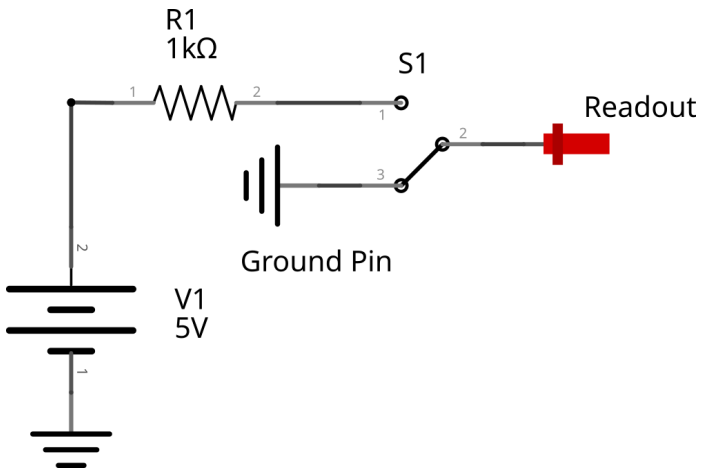


FIGURE A.12 Schematic for connecting a 3-leg switch directly to an Arduino input.

Turn Key and Case Panel Switches

Both the Turn Key and the Case Panel switches of the Sentry unit support either the 2-leg or 3-leg wiring configurations, depending on installer preference and available connectors.

Connection Using Two Legs

When using two legs, prepare two instances of the resistor-cable assembly shown in figure A.11. Connect one leg of the switch to the white signal wire and the other to a ground pin (`LOW`).

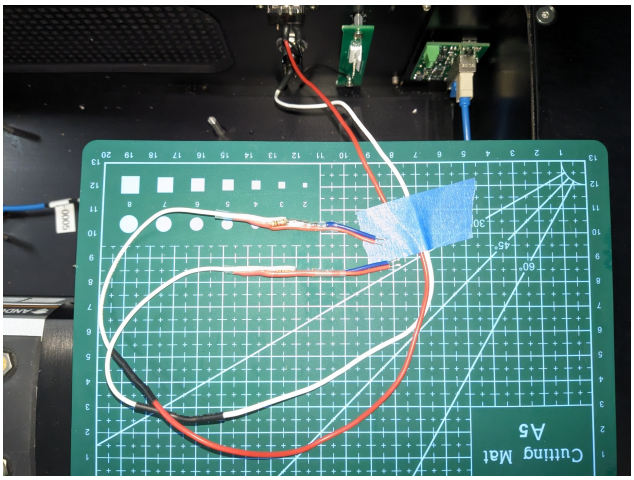


FIGURE A.13 Turn Key or Panel switch wired using the two-leg configuration.

Connection Using Three Legs

In the three-leg configuration, connect the switch to ground, a 5 V output pin, and a digital input pin for probe detection. A 1 kΩ resistor may be placed in series with the 5 V pin for added protection, although this is optional.

Auxiliary Box

The Auxiliary Box (AuxBox) serves as the external interface for status indication and emergency control. It houses an RGB LED display and a physical emergency stop button. This section details the disassembly of the existing AuxBox, mechanical modifications, and complete wiring and reassembly procedures.

Required Materials

- Standard plastic enclosure (existing AuxBox)
- Phoenix Contact circular cable connector (panel mount and cable mount)
- 3× LED indicator (common cathode)
- Emergency stop push-button
- 16 mm drill bit
- 8-wire ribbon cable with keyed connector
- Heat shrink tubing, wire stripper, soldering tools

Disassembly and Mechanical Preparation

Remove all existing components from the original AuxBox. This results in a clean housing ready for modification, as shown in figure A.14.

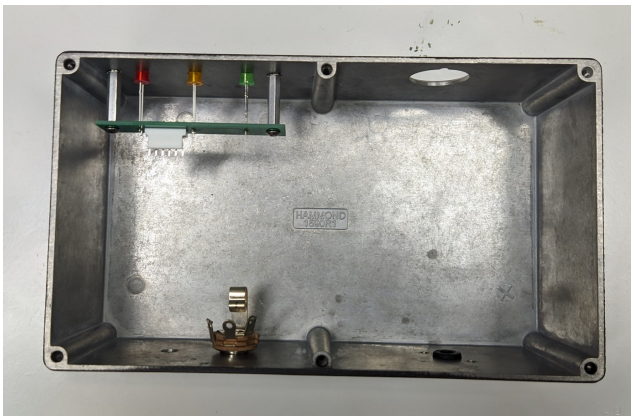


FIGURE A.14 Auxiliary box interior following removal of original components.

Drilling the M16 Connector Hole

Drill a 16 mm hole in the enclosure for mounting the Phoenix Contact panel connector, as shown in figure A.15.

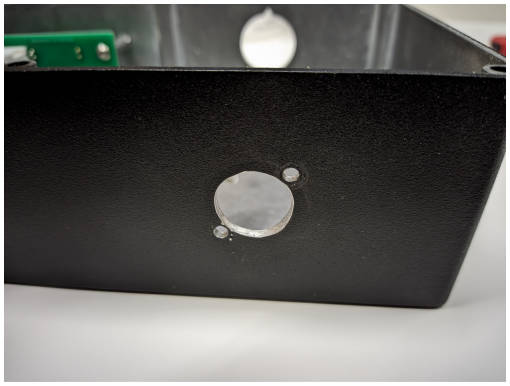


FIGURE A.15 M16 hole drilled into AuxBox for Phoenix Contact panel connector.

Cable Connector Installation

Cutting the cable to a length of approximately 150 mm (6 in) simplifies internal routing and strain relief during assembly.

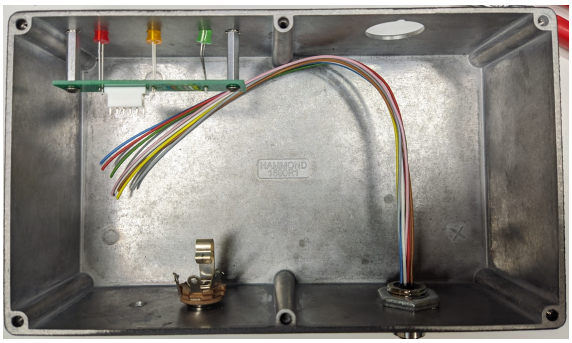
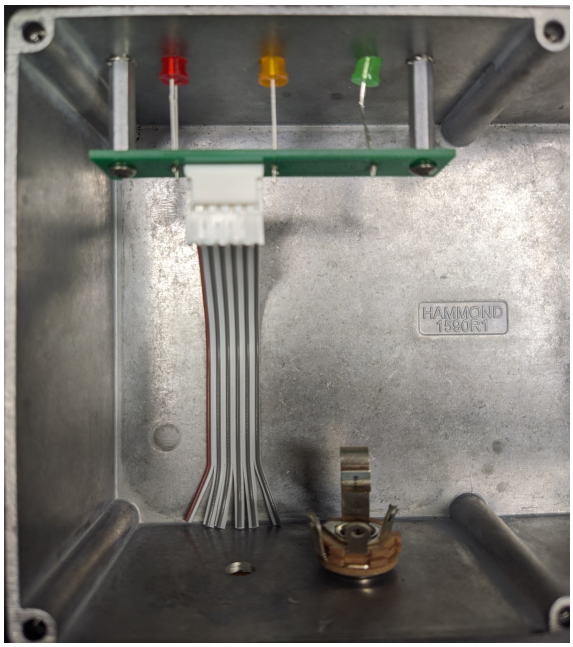


FIGURE A.16 Phoenix Contact cable connector prepared for mounting in the AuxBox.

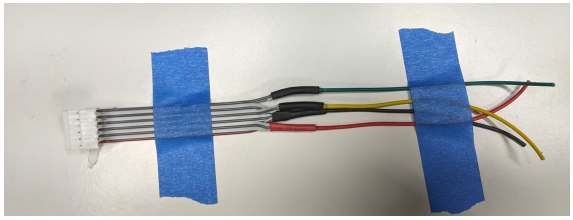
Component Preparation

LED Wires

Ensure that the ribbon cable is oriented such that the elbow of the white keyed connector bends downward, and the red stripe is positioned to the left. See figure A.17 for reference.



(a) Correct elbow orientation of LED ribbon connector.



(b) Red stripe aligned to the left of connector.

FIGURE A.17 LED ribbon cable orientation prior to insertion.

Emergency Stop Button

figure A.18 shows the emergency stop button prepared for installation.

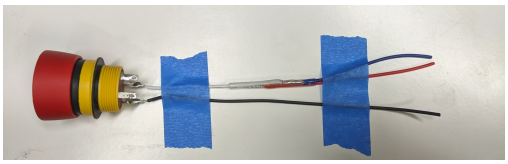


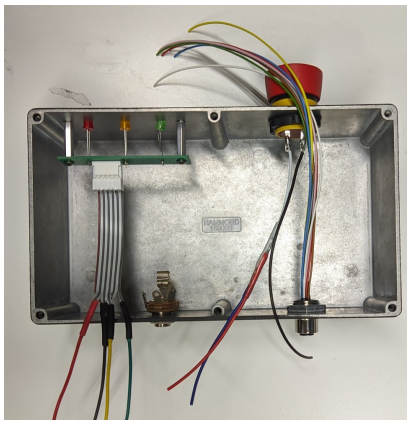
FIGURE A.18 Emergency stop button used in the AuxBox.

Wiring Configuration

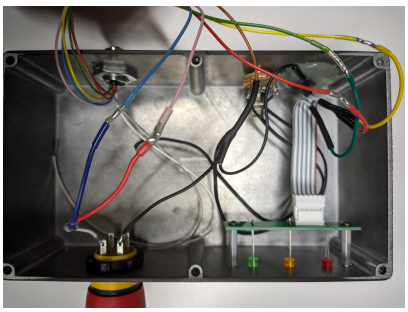
tableau A.2 summarizes the wire color conventions and their corresponding electrical functions within the AuxBox.

Wire Color	**Connection**
Brown	Ground for LEDs and Emergency Stop button
Green	GREEN LED control
Yellow	YELLOW LED control
Red	RED LED control
Pink	3.3 V (`HIGH`) for Emergency Stop button
Blue	Probe (`digitalRead`) for Emergency Stop button
Grey	Not connected
White	Not connected

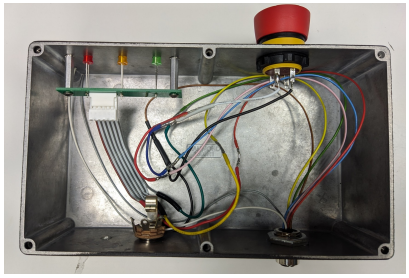
TABLE A.2 Auxiliary box wire color mapping and corresponding signals.



(a) Before wiring.



(b) Completed internal wiring connection.

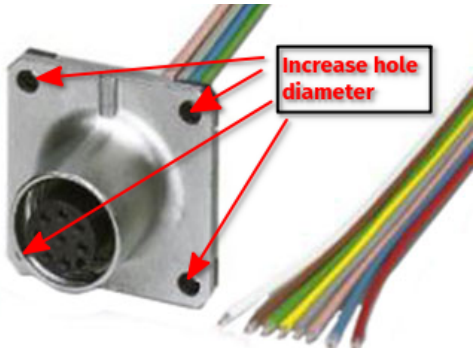


(c) Final assembly and cable management showing all connected components.

FIGURE A.19 Auxiliary box wiring and final integration.

Mounting the Panel Connector to the Sentry Unit

The mounting holes on the square flange of the Phoenix Contact circular panel connector do not align with the threaded holes on the Sentry unit's enclosure. To resolve this, enlarge the flange holes slightly using an over-sized drill bit. Figure A.20 illustrates the panel connector holes to be enlarged, and the connector mounted on the unit.



(a) Square plate holes enlarged to accommodate non-standard alignment.

(b) Panel connector mounted to the rear of the Sentry unit.

FIGURE A.20 Final installation of AuxBox cable connector into the Sentry unit enclosure.

Cable Management and Cleanup

Upon completion of the hardware integration, all internal cables should be carefully routed and secured to ensure system reliability, airflow, and ease of maintenance. Particular attention should be paid to :

— Minimizing cable slack to avoid entanglement or obstruction.
— Grouping cables by function (e.g., power, data, control) using cable ties or adhesive mounts.
— Ensuring that no wires are under mechanical stress or tension.
— Preventing interference by separating low-voltage digital lines from high-current power lines.
— Adding ferrite core to USB cables to prevent electronic interference

figure A.21 shows the final internal layout of the upgraded Lumed Raman System, demonstrating proper cable bundling and separation.

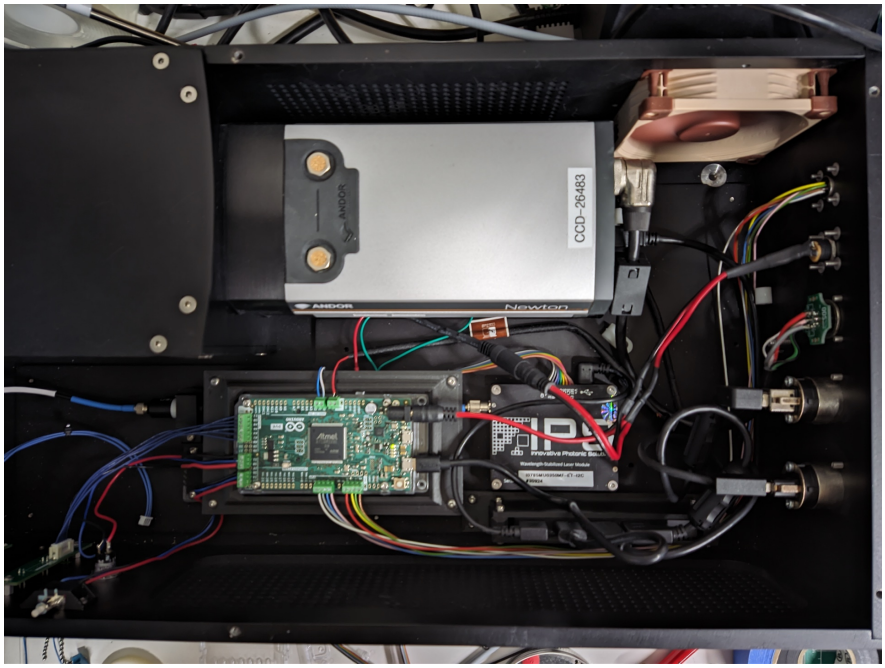


FIGURE A.21 Final cable management layout following full hardware assembly.

APPENDIX B INSTALLATION GUIDE FOR THE LRS SOFTWARE PLATFORM AND ORAS

ORAS has been tested exclusively on Linux-based systems. It is strongly recommended to use *Linux Mint*—specifically the Cinnamon Edition—for optimal compatibility and user experience. [1]

Installation of Linux Mint

Creating a Bootable USB Drive

1. Download the ISO image for Linux Mint : `https://www.linuxmint.com/download.php`
2. Create a bootable USB drive using one of the following tools :
   (a) `https://rufus.ie/en/`
   (b) `https://etcher.balena.io/`
   (c) `https://www.ventoy.net/en/index.html`

Note – Bootable USB Option

The preferred method is to use Ventoy. A 250 GB portable USB SSD with Ventoy and multiple ISO images can streamline future installations. This guide assumes this setup.

—GS

Clean Installation of Linux Mint

Booting the ISO

1. Start from a powered-off PC or laptop.
2. Insert the bootable USB device.
3. Power on the machine and access the BIOS (typically by pressing a key such as `F2`, `Del`, or `Esc` during startup).
4. Disable *Secure Boot* in the BIOS settings.
5. Save changes and reboot.

1. `https://www.linuxmint.com/`

6. Access the boot manager and select the Ventoy USB device.
7. Select the Linux Mint ISO from the Ventoy menu.

Installation Steps

1. Launch the Linux Mint installer from the desktop.
2. Select a language (e.g., English).
3. Choose the keyboard layout (e.g., English (US)).
4. If prompted, do **not** connect to the internet. (it makes installation take longer by updating packages now. We will update the packages later.)
5. Accept multimedia codec installation if prompted.
6. Unmount any mounted partitions if asked.
7. Choose installation type : **Erase disk and install Linux Mint**.
8. Review and confirm partition formatting options.
9. Select the time zone (e.g., Toronto).
10. Choose a username and password.
11. Disable home folder encryption.
12. Begin installation and restart once complete.
13. Remove the USB drive when prompted.

Note – Example Configuration (Asus VivoBook)

— **Name :** user

— **Computer name :** AsusVivoBook

— **Username :** user

— **Password :** password (DO NOT USE "password" as a password...)

First Boot and System Configuration

Installing Updates

1. Connect the system to the internet.
2. Update and upgrade the system :

```
sudo apt update
sudo apt upgrade
```

3. Reboot the system.

Installing Python and Dependencies

Install `pip` :

```
sudo apt install python3-pip
```

Place the ORAS `.whl` file and the Andor driver archive `.tar.gz` in `~/Documents`. Then extract and install the Andor driver :

```
cd ~/Documents/andor
./install_andor
```

Warning – Andor Driver Dependency

According to Andor's `FAQ033.pdf`, the driver requires `libusb`. Install it with :

```
sudo apt-get install libusb-dev
```

Fixing USB Permissions on Linux

Refer to : Techoverflow guide to fix usb permissions

```
wget https://techoverflow.net/scripts/udev-install-usbusers.sh
chmod +x ./udev-install-usbusers.sh
sudo ./udev-install-usbusers.sh
```

Add the user to the `dialout` group :

```
sudo usermod -a -G dialout user
```

Installing ORAS

Install the ORAS package from the `.whl` file :

```
cd ~/Documents/
pip install oras-XXXXX-py3-none-any
```

Creating a Desktop Shortcut

Create a desktop launcher :

```
nano ~/Desktop/ORAS.desktop
```

Insert the following content :

```
[Desktop Entry]
Type=Application
Terminal=true
Exec=/home/user/python/oras/bin/python3 -m oras
Name=ORAS
Icon=accessories-camera
```

Save and close the file using `Ctrl+O`, `Enter`, and `Ctrl+X`.

APPENDIX C FUNCTION REFERENCE FOR THE LRS ARDUINO SAFETY BOARD

TABLE C.1 – Descriptions of functions implemented in the source code of the LRS safety board

Function	**Input (type)**	**Outputs (type)**	**Description**
`setup()`	`void`	`void`	the Arduino setup function (called once at device start)
`loop()`	`void`	`void`	the Arduino loop function (called continuously after setup)
`setupIPSPins()`	`void`	`void`	setup of the IPS laser pins
`setupSwitchesPins()`	`void`	`void`	setup of turnkey and case panel switch pins
`setupLEDPins()`	`void`	`void`	setup of the front panel LED pins
`setupAuxboxPins()`	`void`	`void`	setup of the aux box pins
`getLaserEmission()`	`void`	`bool`	returns `True` if the laser is emitting light (by reading the PD+ pin voltage with `measureVoltageWithHampelFilter()`), `False` otherwise
`getCaseClosed()`	`void`	`bool`	returns `True` if the unit case is closed, `False` otherwise
`getTurnkeyON()`	`void`	`bool`	returns `True` if the turnkey is in the ON position, `False` otherwise
`getAuxButtonSafe()`	`void`	`bool`	returns `True` if the auxbox emergency button has not been pressed, `False` otherwise
`setLaserEnable()`	`bool`	`void`	sets the laser voltage at the laser enable pin based on the input (`True` or `False`)
`setLEDs()`	`int`	`void`	sets the front panel LEDs to the appropriate configuration based on the system state (`0`, `1` or `2`)

Continued on next page

Continued from previous page

Function	Input (type)	Outputs (type)	Description
`setAuxboxLEDs()`	`int`	`void`	sets the auxbox LEDs to the appropriate configuration based on the system state (`0`, `1` or `2`)
`measureVoltagesWithOutlierFilter()`	`int`	`float`	measures the voltage of a pin (pin number as input) with a custom outlier filter and returns the result
`measureVoltageWithHampelFilter()`	`int`	`float`	measures the voltage of a pin (pin number as input) with an implementation of the Hampel outlier detection filter [159, 160]
`printArray()`	`int array[]`, `int size`	`void`	sends the content from an array of numbers via serial communication to a connected computer
`printSystemState()`	`int`	`void`	sends the details of the system state (switch, turnkey, button, laser,...) via serial communication to the connected computer
`insertionSort()`	`int *array[]`, `int size`	`void`	custom implementation of the insertion sort algorithm that sorts an array (input is the array pointer)
`computeMedian()`	`int data[]`, `int size`	`float`	computes the median of the input `data` list of specified `size`
`computeMad()`	`int data[]`, `int size`, `float median`	`float`	computes the median absolute deviation (MAD) of the input `data` list of specified `size` and `median`
`computeMean()`	`int data[]`, `int size`	`float`	computes the mean of the input `data` list of specified `size`
`computeStd()`	`int data[]`, `int size`, `float mean`	`float`	computes the standard deviation of the input `data` list of specified `size` and `mean`