Medical Diagnostics with Surface Enhanced Raman Scattering

Robert Hunter

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Doctor of Philosophy (Ph. D)
degree in Biomedical Engineering

Ottawa-Carleton Institute of Biomedical Engineering
Faculty of Engineering
University of Ottawa

Executive Summary

Raman spectroscopy is a powerful molecular fingerprinting method which measures the vibrational modes of molecules to identify and quantify chemical species. In biomedical spectroscopy, where samples are usually complex mixtures of many molecules, Raman spectra give a biochemical “portrait” that can be used to discriminate between distinct samples. One major technical challenge in implementing Raman spectrometer sensors is the technique’s low intrinsic signal to noise ratio. To amplify the Raman signal, a number of different approaches can be applied. In this thesis two techniques are used; surface enhanced Raman scattering (SERS) from metal nanoparticles along with light-matter interaction enhancement from co-coupling light and sample to a liquid core waveguide.

In order to process the complex spectral data arising from these sensors, a robust signal processing method is required. To this end, we have developed and validated a machine learning spectral analysis platform based on genetically optimized support vector machines (GA-SVM). This work is the subject of Chapter 3. We found that the GA-SVM significantly outperformed the standard statistical based modelling approach, partial least squared, in regression tasks for several different biomedical Raman applications. Furthermore, we found that the use of more complex kernel functions in the SVM yielded superior results. The genetic optimization algorithm was necessary to use these more complex kernel functions because its computation time scales linearly with complexity, whereas the standard brute force approach scales exponentially.

Chapter 4 concerns the development of a Raman sensor used to quantify and identify pathogenic bacteria. This device centres on a microfluidic flow cell which forces bacteria to flow through a hollow-core photonic crystal fiber (HC-PCF) to which the Raman excitation laser is also coupled. The bacteria are also mixed with silver nanoparticles to simultaneously achieve SERS and light-matter interaction enhancement in the sensor. Overall, the fiber and nanoparticles yield a bulk enhancement of 400x for the Raman spectrum. Bacteria are quantified in this system by counting the number of “spectral

events” that occur as cells flow through the HC-PCF in a 15-minute window. This approach achieved very high linearity, as well as an average detection limit of 3.7 CFU/mL. In addition, bacteria are identified by using the same GA-SVM algorithm developed in the preceding chapter. These machine learning models achieved a discrimination accuracy of ~92% when comparing the spectra of the bacteria *S. aureus, P. aeruginosa,* and *E. coli*. In mixed samples of bacteria, the error of quantification increased significantly to 13.3 CFU/mL, but the output of the sensor was highly correlated with the ground-truth bacterial load.

In Chapter 5 we outline the development of a diagnostic scheme for chemoresistance in ovarian cancer based on SERS measurements from cysteine-capped gold nanoparticles. Resistance to chemotherapy was determined based on three factors: the concentration of tumor derived exosomes, the chemical composition of the exosomes, and the concentration of exosome-derived cisplatin. Cisplatin is the drug of interest for this problem, as it is the most basic chemotherapy agent. The system works by first incubating the gold nanoparticles with tumor derived exosomes. The cisplatin therein causes the particles to destabilize slightly, resulting in the aggregation rate of the nanoparticles being proportional to the drug concentration. At steady state aggregation, the magnitude of the Raman spectrum is proportional to the exosome concentration, and the spectrum contains its chemical identity. Using *in vitro* cancer cell lines, we found that resistant cells tend to produce more exosomes and excrete a higher concentration of cisplatin within them. Overall, this sensor exhibited good diagnostic power for chemoresistance particularly in the most common subtype in ovarian cancer.

Acknowledgements

Completion of this thesis would not have been possible without the numerous sources of support and guidance I was given over these past few years. Firstly, I would like to thank my supervisor Dr. Hanan Anis whose guidance, knowledge, and motivation were critical in my development as a researcher. They provided not only the means to conduct this research, but also numerous teaching and professional development opportunities. I also thank all of my other colleagues in the Biophotonics Lab at the University of Ottawa for their support. I am grateful to Drs. Ali Momenpour and Altaf Khetani for providing me training in the lab when I was a new student. Our other teammates Mahamaya Deb, Aseel Mandour, Khaled Atia, Isaiah Brine, Niko Yao, and Vincent Berthiaume all provided valuable help with running, designing, and analysing experiments.

Next, I would like to thank the numerous collaborators with whom I have worked on these projects. It was a privilege to work with so many brilliant individuals from different fields whose diverse expertise allowed me to learn a wealth of new information. Specifically, Drs. Emilio I. Alarcon and Marcelo Muñoz of the BeATS laboratory in the Ottawa Heart Institute provided invaluable help in bringing the SERS flow cytometry system to life. Drs. Alarcon and Muñoz shared their knowledge of bacteria, biofilms, and bio-nanomaterials through many enlightening conversations. Zohra Khatoon, of the same research group, also has my gratitude for running the many bacteria cultures needed to test the system. Next, I would like to thank Drs. Hesham Abdelbary, Paul Beaulé, and Mariam Taha of the orthopaedics department in the Ottawa Hospital who gifted their expertise in clinical practice and medical research which laid the foundation for the joint-infection application of my research. Last but certainly not least I thank Drs. Benjamin K. Tsang and Meshach Asere-Werehene for providing their expertise in ovarian cancer to guide the experiments on chemoresistance diagnosis. It was from their research that the initial idea behind the diagnostic scheme was developed, and their spirit for discovery and innovation drove numerous studies that I would never have guessed I would be a part of.

I would also like to acknowledge my committee members Dr. Jacques Albert, Dr. Pierre Berini, Dr. Jean-Michel Ménard, and Dr. Jürgen Popp for their time, an engaging

thesis defense, and their constructive feedback on my work. I am also grateful to NSERC for their financial support.

Finally, I give my deepest thanks to the friends and family who have supported me throughout this process. Without their love and comradery, I would never have been able to complete this work. I extend special thanks to my partner Amelia French for her incredible patience, emotional support, and keen eye for editing.

Table of Contents

Table of Figures

Table of Tables

Glossary of Terms

AgNP: Silver nanoparticles

ATCC: American type culture collection

AuNP: Gold nanoparticles

CARS: Coherent anti-Stokes Raman scattering

CCD: Charge-coupled device

CDDP: Cisplatin (cis-diamminedichloroplatinum(II))

CFU: Colony forming units

DA: Discriminant analysis

DCM: Dichroic mirror

DIW: Deionized water

DNA: Deoxyribonucleic acid

DOS: Density of states

EIS: Electrical impedance spectroscopy

ELISA: Enzyme linked immunosorbent assay

EM: Endometrioid

EtOH: Ethanol

FBS: Fetal bovine serum

MVA: Multi-variant analysis

MR(S)SA: Methicillin resistant (susceptible) *S. aureus*

NA: Numerical aperture

NIR: Near Infrared

NP: Nanoparticle

OVA: One versus all

OVCA: Ovarian Cancer

OVO: One versus one

OX: Over expression

PCF: Photonic crystal fiber

PCR: Polymerase chain reaction

PDMS: Polydimethylsiloxane

PEEK: polyether ether ketone

PEG: Polyethylene glycol

pGSN: Plasma gelsolin

PJI: Prosthetic joint infection

PLS: Partial least square

FN: False negative

FOM: Figure of merit

FP: False positive

FVU: Fraction of variance unexplained

GA: Genetic algorithm

GABA: gamma-aminobutyric acid

HA: Hyaluronic acid

HAI: Healthcare associated infection

HC-PCF: Hollow core photonic crystal fiber

HGS: High-grade serous

HIV: human immunodeficiency virus

HPLC: High performance liquid chromatography

IAI: Implant associated infection

IR: Infrared

KO: Knock-out

LB: Lysogeny broth

LSPR: Localized surface plasmon resonance

LOD: Limit of detection

PTFE: polytetrafluoroethylene

R6G: Rhodamine 6G

RBF: Radial basis function

RMSE: Root mean square error

RQK: Rational quadratic kernel

SEM: Scanning electron microscopy

SERS: Surface enhance Raman scattering

SMO: Sequential minimal optimization

SPR: Surface plasmon resonance

SRS: Stimulated Raman scattering

SVM: Support vector machine

SVR: Support vector regression

TEM: Transmission electron microscopy

TN: True negative

TP: True positive

UV: Ultraviolet

WT: Wild type

Glossary of Symbols

Photonics

$\boldsymbol{c}$**:** Speed of light (m/s)

$\boldsymbol{D}$**:** Density of states

$\boldsymbol{E}$**:** Electric field (V/m)

$|\boldsymbol{e}\rangle$**:** Excited state

$\boldsymbol{F}$**:** Purcell factor

$|\boldsymbol{f}\rangle$**:** Final state

$\boldsymbol{H}$**:** Magnetic field (A/m)

$|\boldsymbol{i}\rangle$**:** Initial state

$\boldsymbol{k}$**:** Spring constant

$\boldsymbol{m}$ **:** Mass (kg)

$\boldsymbol{m}$**:** Diffraction order

$\boldsymbol{N}$**:** Number

$\boldsymbol{n}$**:** Refractive index

$\boldsymbol{q_k}$**:** Normal atomic coordinate of k^{th} bond

$\boldsymbol{R}$**:** Transition rate (s^{-1})

$\boldsymbol{V}$**:** Volume (m^3)

$\boldsymbol{\alpha}$**:** polarizability (Cm^2V^{-1})

$\boldsymbol{\Gamma}$ $\boldsymbol{or}$ $\boldsymbol{\gamma}$**:** Damping constant

$\boldsymbol{\varepsilon}$**:** Permittivity (F/m)

$\boldsymbol{\lambda}$**:** Wavelength (m)

$\boldsymbol{\mu}$**:** Dipole moment (Cm)

$\boldsymbol{\nu}$**:** Wavenumber (cm^{-1})

$\boldsymbol{\Phi}$**:** Static field (V)

$\boldsymbol{\phi}$**:** Phase

$\boldsymbol{\chi}$**:** Electric susceptibility

$\boldsymbol{\omega}$**:** Frequency (rad/s)

Multivariate Analysis

$\boldsymbol{b}$: Offset

$\boldsymbol{C}$: Cost, box constraint

$\boldsymbol{c_0}$: Polynomial offset

$\boldsymbol{l}$: Length scale

$\boldsymbol{p}$: Polynomial order

$\boldsymbol{\Re}$: Risk

$\boldsymbol{R}$: Kernel radius

$\boldsymbol{w}$: Weight vector

$\boldsymbol{x}$: Input/ predictor

$\boldsymbol{y}$: Output/ prediction

$\boldsymbol{\alpha}$: Lagrange multiplier

$\boldsymbol{\gamma}$: Polynomial scaling

$\boldsymbol{\epsilon}$: Allowed regression error

$\boldsymbol{\kappa}$: Kernel function

$\boldsymbol{\xi}$: Slackness

$\boldsymbol{\sigma}$: Standard deviation

Other

$\boldsymbol{cts}$: Photon counts

$\boldsymbol{j = \sqrt{1}}$: Imaginary number

$\boldsymbol{M}$: Molarity (mol/L)

$\boldsymbol{R^2}$: Pearson's R coefficient

$\boldsymbol{r}$: Radius (m)

$\boldsymbol{t}$: Time (s)

$\boldsymbol{w/v}$: Weight per volume, concentration as percentage

$\boldsymbol{\theta}$: Angle (rad)

Chapter 1 Thesis Objectives and Contributions

1.1 Thesis objectives and outline

A good sensor is comprised of a number of elements that must work together in order to make high quality measurements. These include the transduction element, signal conditioning and amplification, the data processing, and communication. For a system to be useful outside of a research setting, other aspects such as manufacturability, usability, and economic considerations must also be built into the total design. The broad goal of this thesis is to design medical diagnostic systems which synthesize all of these elements, in order to create point-of-care diagnostic systems with real world utility.

Raman spectroscopy is the chosen modality for chemical sensing used in this thesis, due to this technique's numerous strengths. Primary among these is the fact that Raman is a vibrational fingerprint technique, that uses a single incident wavelength to probe the unique normal vibrational modes of a molecule. The resultant spectra can then be applied to quantify chemical species in a sample, or differentiate chemically distinct samples, without the need for label molecules, molecular probes, or biorecognition elements. This contrasts with many conventional assays used in medicine or biological research which rely on expensive reagents that are consumed during the reaction. ELISA assays are an ubiquitous example of this. These techniques are often used to quantify a biological target by capturing it with an antibody, which is then attached to an enzyme like horseradish peroxidase that catalyzes a colour change in a solution, which is quantified via optical absorbance measurements. To isolate targets from a sample before antibody-based detection it is also common to employ Western blots, wherein proteins are separated via gel electrophoresis. This process requires the gel itself, as well as a number of buffer solutions that are consumed in the process. Through designing novel label free chemical sensors, the goal is to reduce the time and costs associated with medical diagnostics, while also improving sensitivity, specificity, and detection limit. Furthermore, when using

a molecular fingerprinting approach such as Raman, it is possible to develop an inherently multiplexed sensor that employs intrinsic chemical differences to identify or quantify multiple desired targets simultaneously. In most other techniques, including other biosensors, multiple biorecognition elements are required in order to develop a multiplex sensor which increases the cost and complexity of manufacturing.

The aforementioned strengths of Raman spectroscopy are hindered in practice due to the fact that Raman scattering cross-sections are very low, meaning the signal-to-noise ratio of these measurements can be poor. A significant body of research on Raman sensors thus relates to methods of signal amplification. The two approaches employed in this thesis are surface enhanced Raman scattering (SERS) and fiber enhanced Raman. The former approach relies on the resonant properties of metal nanostructures to concentrate the excitation field into sub-wavelength sized modes and enhance photoluminescence via the Purcell effect. The latter involves extending the depth of focus of the excitation laser by co-coupling both light and fluid into the same cavity. The theoretic basis of these enhancement techniques, and Raman scattering in general, is provided in Chapter 2.

To successfully employ Raman spectroscopy in multiplex detection, or even single analyte detection in a complex background matrix, advanced multi-variate signal analysis algorithms are needed. For this reason, in Chapter 3 we outline the development of a machine learning approach for Raman spectral analysis based on support vector machines with genetically optimized similarity functions (kernels). The resultant algorithm supports all subsequent sensor development in this thesis as the primary data processing modality.

Sensor development in this thesis was guided by two medical applications; diagnosis of hospital acquired infections, and determination of chemoresistance in ovarian cancer. With respect to infections, the current gold-standard for diagnosis involves culturing of bacteria acquired from patient samples. The primary issue of this technique is the long lag times which result from the rate at which bacteria grow. In some cases, infecting pathogens may not grow under standard culture conditions at all. As such, we sought to

eliminate the need for laborious culturing steps by developing a highly sensitive means of detecting bacteria via SERS in a fiber-based flow cell. This work is the subject of Chapter 4. In our other application, the current approaches used to determine susceptibility to chemotherapy in ovarian cancers are lacking in accuracy and timeliness. Our collaborators at the Ottawa Hospital Research Institute have recently identified a number of biomarkers that could be used to determine chemoresistance. One of these potential markers is exosome-derived chemotherapy drug, for which we developed a sensor based on a kinetic SERS assay on amino acid functionalized gold nanoparticles. This work is outlined in Chapter 5.

The experiments conducted over the course of this thesis also provide the basis for future research directions. These include bio-functional SERS substrates, novel microfluidic systems, and advanced microscopy techniques. These ideas are elaborated upon in Chapter 6 with the goal of providing new researchers a starting point for ideating future projects.

1.2 Novelty and contributions

The work conducted in this thesis has produced a number of innovations and new scientific data that are of interest to the fields of chemical sensing, chemometrics, Raman spectroscopy, and medicine. These include:

- The use of complex hybrid kernel functions in SVMs to analyse Raman spectra, supported by a genetic algorithm whose analytical fitness function greatly reduces processing times compared to other techniques. In the past, only simple kernel functions have been used, due in part to the computational load required to optimize more complex functions. Here, we have shown that complex hybrid kernel functions yield superior models in many Raman chemometric problems.
    - This was the first work to our knowledge to apply an analytical fitness function in the optimization of SVM regression, thereby facilitating the use of complex kernels with minimal computation time.

- Building our group's previous work on HC-PCF Raman, we developed the first SERS flow cytometry system using an optical fiber based flow cell, which supports both SERS and depth-of-field enhancement of Raman spectra. This is applied to the rapid and culture free identification and quantification of bacteria in serum of synovial fluid, achieving a low detection limit in a short assay time. Eliminating the need to culture bacteria in order to get an accurate diagnosis has huge potential for improving quality of care.

- Designed a novel sensing system for the chemotherapy drug cisplatin and tumor derived exosomes. Cisplatin was found to modulate the stability of cysteine-capped gold nanoparticles, yielding an aggregation rate that was proportional to the drug concentration. As the particles stochastically aggregated about exosomes in the sample, the intrinsic SERS spectra generated thereof could be used to differentiate different subtypes of ovarian cancers as well as classify their chemosusceptibility.
    - Developed a simple means of synthesizing highly stable cysteine capped gold nanoparticles.
    - Used this sensor to analyse the effect of plasma gelsolin, a protein suspected of playing a role in chemoresistance. Thus the sensor was found to be useful both as a diagnostic, and research tool.

1.2.1 Journal Articles

1.2.2 Produced as part of this thesis

Robert A. Hunter, Meshach Asare-Werehene, Aseel Mandour, Benjamin K. Tsang, Hanan Anis. (2022) "Determination of Chemoresistance in Ovarian Cancer by Simultaneous Quantification of Exosomes and Exosomal Cisplatin with Surface Enhanced Raman Scattering." Sensors and Actuators B: Chemical. doi:10.1016/j.snb.2021.131237

Robert Hunter, Ali Najafi Sohi, Zohra Khatoon, Vincent R. Berthiaume, Emilio I. Alarcon, Michel Godin, Hanan Anis. (2019) "Optofluidic SERS platform for rapid bacteria detection in serum." Sensors and Actuators B: Chemical. doi:10.1016/j.snb.2019.126907

Robert Hunter, Hanan Anis. (2018). “Genetic support vector machines as powerful tools for the analysis of biomedical Raman spectra.” Journal of Raman Spectroscopy. doi:10.1002/jrs.5410

1.2.3 Related to this thesis

Mahamaya Deb, **Robert A. Hunter**, Mariam Taha, Hesham Abdelbary, Hanan Anis. (2022) “Rapid detection of bacteria using gold nanoparticles in SERS with three different capping agents: thioglucose, polyvinylpyrrolidone, and citrate” in review

1.2.4 As collaborator

Marcelo Munoz, Antony El-khouri, Cagla Eren Cimenci, Mayte Gonzalez, **Robert Hunter**, David Lonboni, Fabio Variola, Liane M. Rossi, Ana Maria Edwards, Erik J. Suuronen, Emilio I. Alarcon. (2021) “Riboflavin surface modification of Poly Vinyl Chloride for light triggered control of bacterial biofilm and virus deactivation.” ACS Applied Materials and Interfaces. doi: 10.1021/acsami.1c08042

Chapter 2 Introduction to Raman Spectroscopy and Enhancement

2.1 Introduction

This chapter will provide a brief introduction to Raman scattering and its associated techniques. Particular focus will be placed on methods by which inelastic scattering can be enhanced so that this approach can be applied to the detection of dilute analytes.

2.2 Raman Scattering

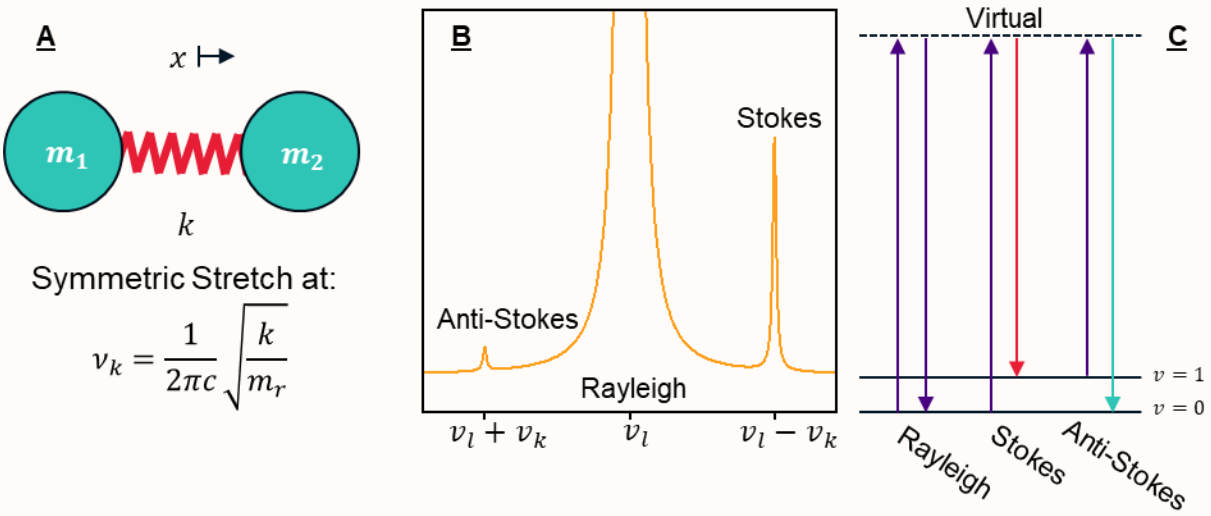

Figure 1: (A) Example of a diatomic molecule considered as two masses (atoms) and a spring (bond), (B) and the resulting scattering at Rayleigh, Stokes, and anti-Stokes Raman frequencies. (C) The same as visualized using quantum energy levels.

Raman spectroscopy is a vibrational spectroscopy technique which uses a single excitation wavelength to probe the energy distribution of bond vibrational modes in a molecule. This technique has been applied in numerous chemical analyses from pharmaceutical quality control to medical diagnostics. Raman scattering arises from the inelastic scattering of photons from molecules, wherein energy is exchanged between a bond vibration and the scattered photon. To describe this phenomenon from the perspective of classical mechanics (derivation adapted from [1]), consider a diatomic

molecule as two atoms with masses m_1 and m_2 bound together by a linear spring with a Hooke's law constant of k. Each of these two molecules can move in three dimensions, thus this diatomic molecule has six degrees of freedom. However, three of these are simply translation of the molecule, and two more are rotations, thus we only have one vibration wherein the atoms synchronously stretch and compress the bond without moving the centre of mass. The dynamics of bond length x in this vibration can be described by a 2nd order differential equation which is the sum of the atomic inertia and restoring force of the bond

$$F = \frac{dm_r^2}{d^2x} + kx. \quad (1)$$

In this equation we consider some arbitrary applied force F, and the reduced mass $m_r = \left(\frac{1}{m_1} + \frac{1}{m_2}\right)^{-1}$. Solving for the impulse response of this system results in a resonant vibration whose natural frequency expressed in terms of wavenumber ν_k, is

$$\nu_k = \frac{1}{2\pi c}\sqrt{\frac{k}{m_r}}, \quad (2)$$

where c is the speed of light in vacuum. Even in this simple model, it can be seen that different combinations of atoms and bonds will lead to a unique set of normal vibrational frequencies which, if measured, can be used as a spectral fingerprint of a molecule [2]. In the classical model presented here, this resonant vibrational frequency represents the ground energy level of the discrete quantum model for a harmonic oscillator.

To generalize this analysis to consider molecules with multiple bonds and normal vibrations let us replace x with q_k, which is a normal coordinate for describing the motions of atoms in the k^{th} normal vibration. As a bond vibrates, there is a sinusoidal change in this distance between the two atoms in our diatomic molecule given by

$$q_k = q_k^0 \cos(2\pi c \nu_k t + \phi), \quad (3)$$

where q_k^0 is the vibrational magnitude and ϕ is the vibrational phase, considering $q_k = 0$ as the average bond length. This periodic change in bond length also leads to an associated change in the bond polarizability α. This factor describes the magnitude of the dipole moment (μ) generated in response to an applied electric field as $\mu_k = \alpha(q_k)E$. Even without knowing an analytical description of $\alpha(q_k)$, we can expand it as a Taylor series and retain only the 0th and 1st order dependencies on the bond length to describe the total dipole for normal vibration k as

$$\mu_k = \left(\alpha_0 + \frac{\partial \alpha}{\partial q_k}\Big|_{q_k=0} q_k^0 \cos(2\pi c \nu_k t + \phi)\right) E. \quad (4)$$

If an electromagnetic wave described by $E = E_0 \cos(2\pi c \nu_0 t)$ is incident on this vibrating molecule, the overall oscillating dipole moment then becomes

$$\mu_k = \alpha_0 E_0 \cos(2\pi c \nu_0 t) + \frac{1}{2}\left(\frac{\partial \alpha}{\partial q_k}\right) q_k^0 E_0 [\cos(2\pi c[\nu_0 - \nu_k^{+1}]t + \phi) + \cos(2\pi c[\nu_0 + \nu_k^{-1}]t + \phi)]. \quad (5)$$

Where α_0 is the average polarizability, and $\partial\alpha/\partial q_k$ is the change in polarizability with respect to the vibrational amplitude of the k^{th} bond mode. We observe that the resulting dipole has three components. The first is at the frequency of the incident radiation, causing the molecule to elastically scatter light. This is known as Rayleigh scattering, and accounts for the majority of the radiation scattered from a molecule. The second and third oscillate respectively at the beat frequencies between the excitation frequency and the two normal molecular vibrations one level above and below its current vibrational state. These describe Stokes and anti-Stokes Raman scattering respectively. At room temperature, the majority of molecules will be in their ground vibrational states, therefore Stokes scattering dominates over anti-Stokes under normal conditions.

From this analysis, we can also derive some insight into the radiated intensity of Raman scattering. An electric dipole radiates a total power of

$$P = \frac{8\pi^3 \mu_0^2 c \nu^4}{\varepsilon_0}, \tag{6}$$

where μ_0 is the amplitude of the dipole moment, and ε_0 is the permittivity of free space. For Raman scattering, $\mu_0 = \frac{1}{2}\left(\frac{\partial \alpha}{\partial q_k}\right) q_k^0 E_0$ and $\nu = \nu_0 \pm \nu_k^{+1} \cong \nu_0$ because we assume that the laser frequency is much greater than the vibrational frequency. Thus, we find that Raman scattering intensity is linearly proportional to the laser intensity, and scales with the fourth power of the incident wavenumber ν_0^4.

Additional understanding can be gleaned from the quantum description of Raman scattering. In this picture, the energy of the incident radiation is in between the resonant energy of vibrational and electronic transitions. Therefore, many possible “virtual” states are excited rather than a real state of the molecule. The transition rate describing Raman scattering $R_{|i\rangle \to |f\rangle}$, which is proportional to the scattered power, is given by

$$R_{|i\rangle \to |f\rangle} = \sum_e \left| \frac{\langle f | \mu_{ef} | e \rangle \langle e | \mu_{ie} | i \rangle}{\omega_0 - \omega_{ei} + j\Gamma_e} \right|^2 |E_0|^2 D(\omega_{fi}), \tag{7}$$

where $|i\rangle$, $|e\rangle$, and $|f\rangle$ are the initial, excited (virtual), and final states respectively, μ_{ef}, and μ_{ie} are the perturbations of the Hamiltonian coupling between the subscript states, and Γ_e is the transition damping due to the lifetime of $|e\rangle$ [3]. In agreement with the classical description derived earlier, we find that the rate is proportional to the intensity of the excitation source which has frequency $\omega_0 = 2\pi c \nu_0$. Additionally, the transition rate is also proportional to the density of final states at the frequency difference between $|i\rangle$ and $|f\rangle$ which is given by $D(\omega_{fi})$. In the case of Stokes Raman scattering $\omega_{fi} = \omega_0 - 2\pi c \nu_k^{+1}$. The idea of enhancing Raman scattering by increasing the local $D(\omega_{fi})$ by placing a molecule in a cavity whose resonance is near ω_{fi} is the basis of surface enhanced Raman scattering which is the subject of Section 2.5.1.

By measuring the spectrum of these Raman shifted photons, we can observe the normal vibrational energies of all modes within the molecule that modulate its polarizability, thus acquiring its unique spectral fingerprint. These spectra are usually

shown as a Raman shift, with the frequency expressed in wavenumbers (cm^{-1}) relative to the excitation (Rayleigh) peak at zero. Stokes shifts are considered to be positive Raman shifts in this convention.

2.3 Band Assignment

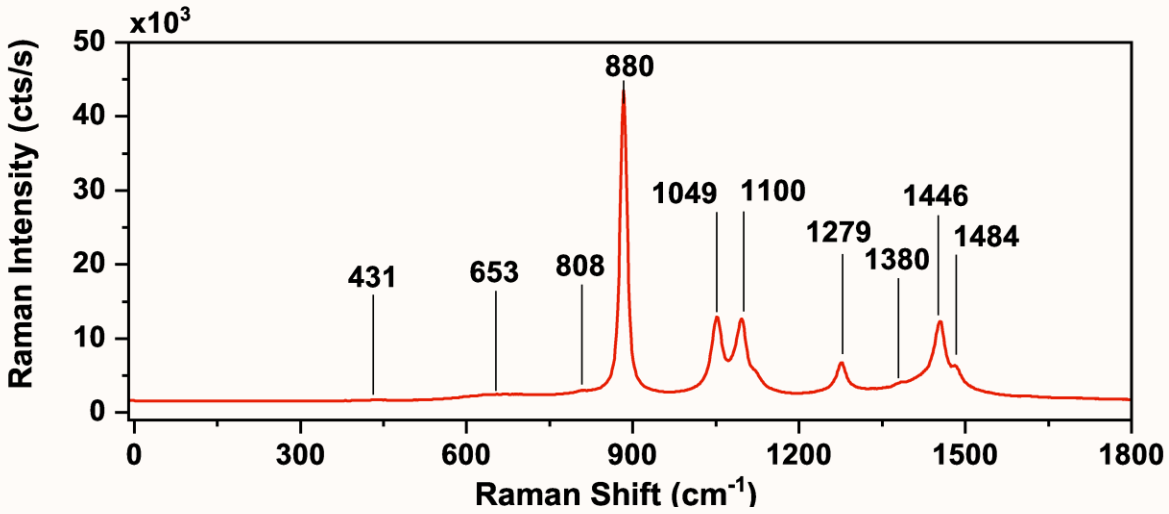

Figure 2: The Raman spectrum of ethanol in the fingerprint region with the peaks labelled with their corresponding Raman shift.

With this basic understanding of Raman scattering, let us now consider the factors which contribute to generating real spectra, and how we can determine the bond vibrations associated with a Raman band. In general, a given molecule with N atoms will have $3N-6$ vibrational degrees of freedom (or $3N-5$ for linear molecules). Each normal vibrational frequency k can be measured as a Raman scattered field so long as its polarizability modulation, $\frac{\partial \alpha}{\partial q_k}$ as seen in equation (5), is not equal to zero. Whether or not a certain vibration is "Raman active" depends on the geometry of the molecule. Specifically, if a molecule has a center of inversion symmetry, then any vibrations that are anti-symmetric with respect to this point will have $\frac{\partial \alpha}{\partial q_k} = 0$. In addition to these normal vibrations, combination terms such as harmonics and beat vibrations are also present, but these are often too weak to be observed.

Take for example the Raman spectrum of ethanol displayed in Figure 2. Ethanol has 9 atoms, and therefore has 21 vibrational degrees of freedom, all of which are theoretically

Raman active because ethanol lacks inversion symmetry. The spectrometer used in this thesis captures Raman spectra in the “fingerprint” region up to 1800 cm^{-1}. The Raman bands of ethanol in Figure 2 are labeled with their corresponding vibrational frequencies and assigned in Table 1. The most straight forward method to determine these band assignments is to simply find tabulated data as I have done in this example [4]. This is relatively easy if the molecular constituent of the sample is known, but may be more complicated in samples containing many molecules and/or those whose identity is not known *a priori*. In these cases, band assignments are considered tentative at best. If the structure of a certain molecule is approximately known, one can also determine the vibrational mode energies using *ab initio* simulations and compare the results to Raman spectra [5].

Table 1: Band assignment for ethanol Raman spectrum [4]

Raman Band (cm^{-1})	Assignment
431	CCO bend
653	COH bend
808	CH_2 rock, CH_3 rock
880	CCO symmetric stretch
1049	CCO asymmetric stretch
1100	CH_3 rock, COH deformation
1279	CH_2 wag
1380	CH_2 wag, CH_3 symmetric deformation
1446	CH_3 asymmetric deformation
1484	CH_2 deformation

There are several other considerations that will effect a Raman measurement that may also need to be accounted for. For example, in liquids Raman bands are broadened due to the damping effect of molecular collisions within the fluid and hydrogen bonding with the solvent [1]. This means that bands regularly overlap, and rotational modes cannot be resolved. In solid crystals, there are also long-range lattice vibrations that will appear in the Raman spectrum.

2.4 Raman Spectral Acquisition

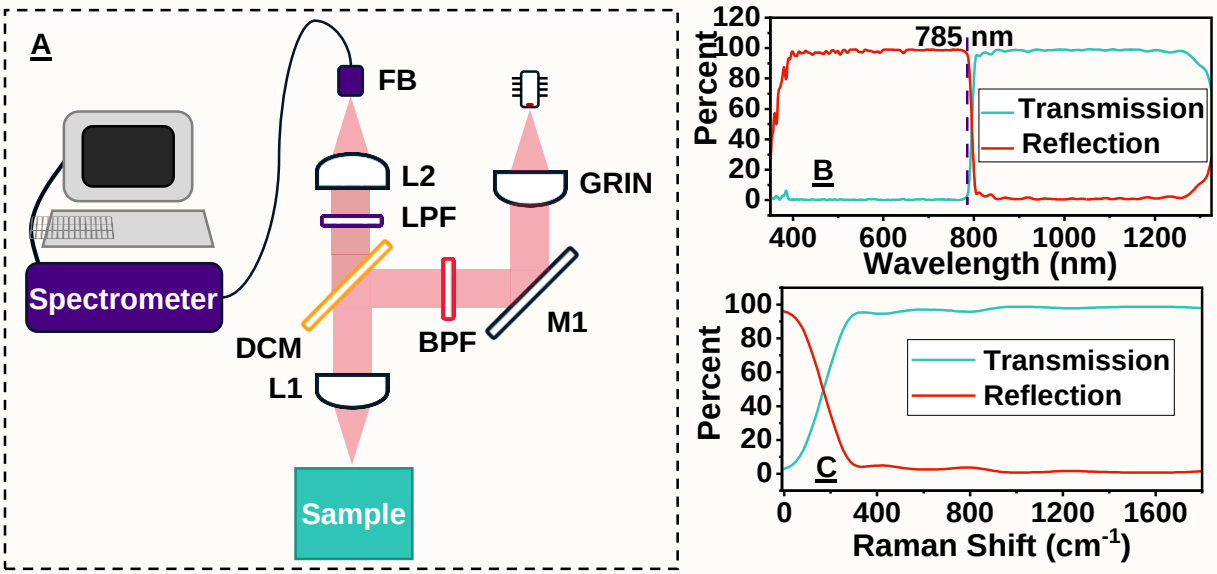

Figure 3: Basics of Raman spectral acquisition. (A) Schematic of a simple Raman spectrometer wherein a laser is collimated by a GRIN lens onto a tuning mirror (M1). The laser is then filtered by a bandpass filter (BPF), and reflected by a longpass dichroic mirror (DCM) before being focused on the sample through a lens (L1). L1 collects Raman scattered photons from the sample, whose longer wavelength allows them to pass through the DCM and a longpass filter (LPF) before being focused through another lens (L2) onto a multimode fiber bundle (FB) into the spectrometer. (B) Transmission and reflection of the DCM used in this thesis, namely DMLP805 from Thorlabs. (C) The DCM properties from (B) in the fingerprint region of the Raman spectrum. All DCM properties were acquired from the Thorlabs website.

The basic set-up for Raman measurements consists of 1) a single frequency source, 2) filtering optics, 3) sample and illumination optics, and 4) the spectrometer. The example in Figure 3A illustrates the basic Raman system used in this thesis. The laser is a 785 nm distributed Bragg reflector (DBR) laser, driven by a current and thermoelectric cooling controller. The DBR laser is a class of single frequency semi-conductor lasers wherein the gain region is flanked by two elements of periodic refractive index (Bragg gratings). The result of the periodic structures is that a certain wavelength, which is phase matched

to the periodicity of the gratings, will be strongly reflected in the cavity due to constructive interference from reflections from each grating element [6]. The result is a laser output with a much narrower spectral line width than provided by the Fabry-Perot cavity or gain spectrum of the active layer alone. This is important in Raman spectroscopy in particular because Raman bands are also spectrally narrow. Since the laser frequency distribution is convolved with the Raman frequency distribution, the use of simple lasers with broad emission bands would significantly reduce the resolution of the spectrometer. DBR lasers are not the only class of single frequency excitation sources that may be used for Raman spectroscopy. Any narrow bandwidth source such as atomic emissions from a gas discharge lamp or distributed feedback, external cavity, or vertical cavity surface emitting lasers can be used.

This laser reflects off a tuning mirror and into a filter cube which consists of an excitation filter (bandpass, BPF), a dichroic mirror (DCM), and an emission (longpass, LPF) filter. The purpose of a filter cube is to pass the excitation radiation to the sample, while passing only the emission from the sample to the detector. The bandpass excitation filter is chosen such that the design wavelength of the filter is the same as the laser, and is used to eliminate any spectral sidebands produced by the laser and further reduce its bandwidth. The DCM can be chosen as either a longpass or shortpass mirror, and this depends on the specific layout of the system. In this example, a longpass mirror is used, thus reflecting the excitation beam towards the sample, and transmitting the Stokes-Raman scattered photons. In order to collect as much of the Raman spectrum as possible, the excitation wavelength should be close to the red edge of a longpass mirror or the blue edge of a short-pass mirror. Thus the Raman spectrum will primarily lie in the transmission or reflection band respectively. The longpass filter located after the DCM is used in order to further reduce the amount of excitation frequency that enters the spectrometer. This is necessary because Rayleigh scattering and reflection dominates over Raman scattering by several orders of magnitude. Again, the excitation wavelength should be as close to the red edge of the reflection band as possible. In the basic Raman system used in this thesis, the DCM is a longpass 805 nm cut-on mirror with a reflection band from 400 – 785 nm and a transmission band from 825 – 1300 nm as shown in Figure 3B. Note that the transition region of the mirrors means that Raman shifts smaller than 350 cm^{-1} will not be

efficiently collected by this system as shown in Figure 3C where the DCM specifications are shown in terms of Raman shift from a 785 nm source.

The beam reflected from the DCM is incident on a lens that focuses the light onto a sample to excite scattering. Since Raman scattering is isotropic, it can be collected in any direction. Thus in the simplest case, the same lens is used to both excite and collect scattering. In our experiments, the lens was a plano-convex aspheric with NA = 0.6 and f = 10 mm. Raman scattered photons collected by this lens pass through the DCM and are coupled into a spectrometer. These systems are usually either grating or Michelson interferometer based. In the case of a grating based system, the spectrometer consists of an array of imaging optics which project the image of the entrance slit onto a camera detector. The imaging array contains an area of infinity space where the grating is located. The light incident on the grating is diffracted into a number of grating orders m, with angular dependence on the incident wavelength λ. The grating equation is given by

$$\sin(\alpha) + \sin(\beta) = \frac{m\lambda}{a}. \qquad (8)$$

This angular deflection of the beam results in a spatial distribution of a spectrum over the camera. In an interferometer based spectrometer, the circular entrance pupil is imaged through a Michelson interferometer onto a single detector. One of the arms of the interferometer is attached to a motor that moves the corresponding mirror. The resulting change in path length changes the interference pattern measured by the detector and the spectrum of the input can then be extracted from the interference pattern by means of Fourier analysis.

Using these basic ingredients for instrumentation, it is possible to develop Raman spectroscopy systems to suit a wide variety of applications. For example, Raman scattering can be excited and collected through a fiber optic probe to perform endoscopy [7], or the exaction and collection optics can be spatially offset to efficiently collect Raman scattered photons from deep inside scattering media [8]. Microscopes can be constructed by adding a means of scanning the laser beam waist over an area of interest, generating a spectrum at each point [3]. Additional instrumentation can be added to isolate Raman

scattering from competing fluorescence, such as by modulating the laser intensity and using a lock-in detector to temporally discriminate these signals [9]. Raman instruments can also be added into other imaging/detection modalities for added functionality. For example, Raman spectra can be acquired from particles trapped in optical tweezers to perform various analyses [10].

2.5 Enhancement of Raman Scattering

Molecular fingerprinting with Raman spectroscopy is potentially very powerful, but this technique has a significant drawback; Raman scattering is very weak, and only about 1 in 10^7 photons are inelastically scattered. The overall measured intensity of Raman scattering is proportional to the incident excitation intensity, the number of scatterers, the square of the polarizability modulation $(\partial\alpha/\partial q_k)^2$, the fourth power of the incident wavenumber ν_0^4, the local photon density of states (DOS) $D(\omega_{fi})$, as well as instrument specific factors such as detector efficiency and losses from optical components. In many cases there are certain limits placed on these parameters, and they cannot be increased beyond a set point. For example, in biomedical Raman spectroscopy the laser power cannot be so high as to damage cells or tissue. Additionally, the laser wavenumber should be in the biological optical window, which resides in the NIR regime, in order to minimize absorbance and background fluorescence. Using these longer wavelengths is necessary, but significantly reduces Raman intensity due to the ν_0^4 dependency. A substantial body of research exists regarding methods to amplify Raman scattering in order to increase its utility. The subsequent sections will describe the two enhancement methods employed in this thesis; surface enhanced Raman scattering, and fiber enhancement. Other methods employed in this field will be briefly mentioned at the end of this section.

2.5.1 Surface Enhanced Raman Scattering (SERS)

One of the most common approaches to Raman amplification is surface enhanced Raman scattering (SERS), which occurs when a molecule is within the localized surface plasmon resonance (LSPR) mode of a metal nanostructure. This phenomenon results from the unique optical properties of metals, due primarily to the free electrons in the conduction band. The Raman scattering enhancement that results from optical excitation

of metal nanostructures is due to the local resonant increase in electromagnetic field intensity, the Purcell effect, charge transfer resonances, and molecular resonances. These elements are in turn dependant on the size, shape, and material(s) constituting the nanostructure. Numerous different structures made of conductive materials such as gold [11], silver [12], or copper [13] have been reported. The most common of these is the metal nanoparticle, which is often formed from a bottom-up synthesis involving reduction of metal salts. The result is an approximately spherical particle which is then made to interact with a molecule of interest in order to acquire SERS spectra. Let us consider a model for the secondary emission enhancement experienced by a molecule near the surface of such a spherical nanoparticle.

2.5.1.1 Theory of SERS

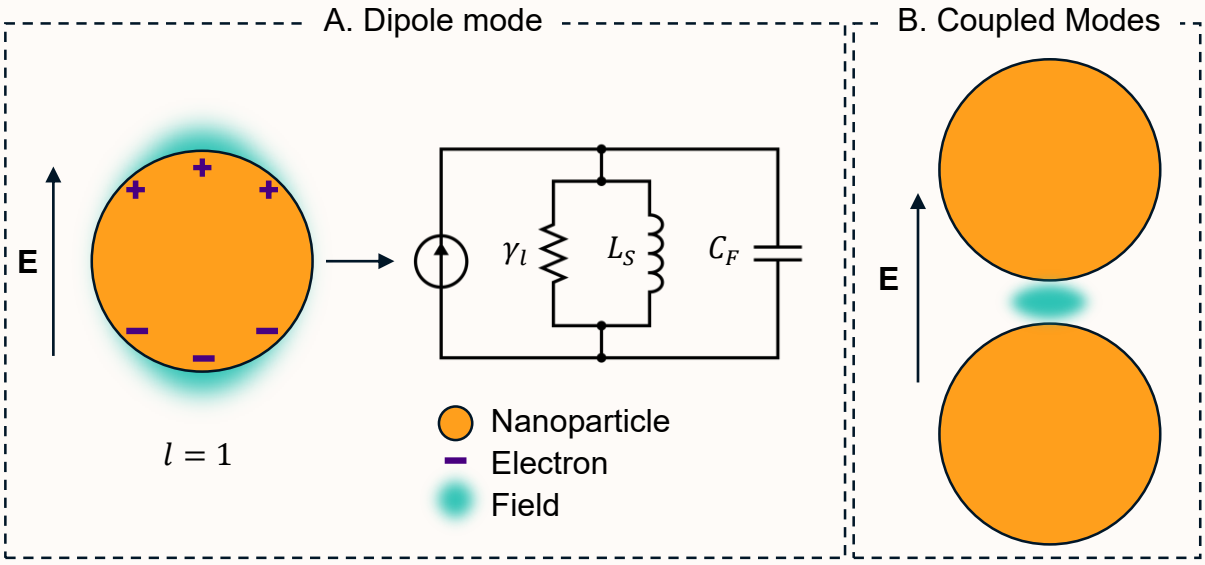

Figure 4: Illustration of metal nanoparticle resonators used in SERS. A) Illustration of the main dipole localized surface plasmon resonance mode of a metal nanoparticle and its associated circuit model. Resonance arises from the exchange of energy between the static field capacitance C_F and the inductance from electron momentum L_S with losses arising from the metal and due to radiation γ_l. B) Qualitative illustration of a coupled mode between two nanoparticles illustrating the plasmonic "hotspot" that forms between them.

A completely unified theory of SERS accounting for electromagnetic as well as resonant electrochemical enhancement effects is still an area of active research and is therefore beyond the scope of this thesis. In this section, a simple model that accounts for the majority of relevant enhancement effects will be shown, as adapted from the following texts [3,14–17] which the interested reader is encouraged to consult for more

detail. If we consider a single metal nano-sphere with radius a in a homogeneous dielectric with permittivity ε_d, and input an electromagnetic field such that $\lambda \gg a$, we can consider the local field in the vicinity of the particle to be homogeneous at any time t and therefore apply an electrostatic approach to determine the modes of the particle that satisfy the Laplace equation

$$\nabla^2 \Phi = 0, \tag{9}$$

where Φ is the electric potential. We can apply the known solutions to this equation, which take the form of Legendre polynomials [18], and utilize the electromagnetic boundary condition requirements to find a solution. Furthermore, consider only a single molecule in the vicinity of the nanoparticle on its polar axis. Thus, for this analysis we may neglect the angular dependency of the polynomials. The boundary conditions are continuity of the electric field tangential to the interface, and continuity of the displacement field normal to the interface. The latter requirement necessitates that the real parts of the permittivities satisfy

$$l\varepsilon_m(\omega_l) + (l+1)\varepsilon_d(\omega_l) = 0 \mid l > 0, \tag{10}$$

where ω_l is the resonant frequency of the l^{th} mode (related to the l^{th} Legendre polynomial). If the permittivity of the metal is described by the Drude model, then

$$\varepsilon_m(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + j\omega\gamma}, \tag{11}$$

where ω_p is the plasma frequency, and γ is the damping rate of the metal. If ε_d is constant over all frequencies then the resonant frequency of the l^{th} mode is that which satisfies equation (10)

$$\omega_l = \sqrt{\frac{l\omega_p^2}{l + (l+1)\varepsilon_d} - \gamma^2} \tag{12}$$

It should be noted that, for completeness, the permittivity in equation (11) treats the metal as a pure plasma and ignores the contribution of the bound valence electrons. A

more accurate model can be obtained by including Lorentz terms in place of the 1 in the right-hand side of equation (11). These are omitted here for the sake of clarity, and their inclusion does not significantly alter the concepts described next.

These LSPR modes supported by the particle have their maximum intensity directly at the surface of the nanoparticle ($r = a$). The $l = 1$ mode takes the form of a dipole resonance parallel to the applied electric field that is set up due to the longitudinal exchange of energy between the electric potential (capacitance), and the motion of conduction electrons in the metal (inductance). The illustration in Figure 4A shows this dipole mode along with an equivalent circuit model based on the work by Engheta *et al* [19]. This LSPR mode has an associated dipole moment, meaning that it is capable of in and out coupling of radiated energy. All higher order modes more closely resemble surface plasmon polariton modes which cannot be directly coupled to free-space radiation.

As shown in equation (7), the Raman decay rate is proportional to the term $|E_0|^2 D(\omega_{fi})$, the product of the electric field magnitude and the local DOS. The ratio of this quantity in the presence and absence of the metal nanoparticle is the overall enhancement, also know as the Purcell factor F

$$F = \frac{|E_D|^2 D_{local}(\omega_{fi})}{|E_0|^2 D(\omega_{fi})}, \qquad (13)$$

where $D_{local}(\omega_{fi})$ is defined as the local DOS in the vicinity of the metal nanoparticle due to the Purcell effect. This effect, originally outlined in 1946 by Edward Purcell, describes the reduction in the spontaneous emission lifetime of a quantum emitter placed in a cavity with allowed modes at the emission frequency [20]. The rate is increased proportionally to the quality (Q) factor of the resonator and inversely proportional to the mode volume. First, let us consider Raman scattering enhancement due to the buildup of local field intensity due to the LSPR modes. Specifically, we need only consider the $l = 1$ mode to which external radiation can couple, and specify that the scattering molecule is on the

polar axis of the nanoparticle. The associated electric field magnitude outside the nanoparticle $|E_D|^2$ is then given by

$$|E_D|^2 = |E_0|^2 \left(\left| 1 + 2\frac{a^3}{r^3} \frac{\varepsilon_D - \varepsilon_m}{\varepsilon_D + 2\varepsilon_m} \right|^2 \right) | \, r > a \;\&\; \theta = 0, \pi. \qquad (14)$$

Thus the field enhancement is

$$\frac{|E_D|^2}{|E_0|^2} = \left| 1 + 2\frac{a^3}{r^3} \frac{\varepsilon_D - \varepsilon_m}{\varepsilon_D + 2\varepsilon_m} \right|^2. \qquad (15)$$

This equation clearly shows that the field enhancement decreases as r^3 as the scattering molecule moves further from the nanoparticle surface. Therefore, field enhancement is maximized for molecules bound directly to the particle.

The local DOS is also enhanced because the molecule in the vicinity of the nanoparticle can radiate energy into the LSPR modes as well as free space modes. Photons can be scattered into any of the LSPR modes, but only the $l = 1$ mode can then couple into free space radiation. Therefore, there is a competing effect between Raman scattering enhancement by coupling scattering into the $l = 1$ versus a quenching effect for scattering coupled into the $l \geq 2$ modes. To develop a Purcell factor for this effect, the Q-factor for an LSPR mode can be estimated from the Lorentzian linewidth f_{LSPR} of its frequency response. This is given by

$$f_{LSPR}(\omega_{fi}) = \frac{1}{2\pi} \frac{\gamma_l}{\left(\omega_{fi} - \omega_l\right)^2 + \gamma_l/4}, \qquad (16)$$

where γ_l is the damping rate of the l^{th} mode which is equal to the metal damping γ for all $l \geq 2$ modes and the sum of metal damping and radiation decay γ_{rad} for the dipole mode. The effective LSPR mode volume is

$$V_{LSPR,l} = \frac{4\pi a^3}{(l+1)^2 \varepsilon_d}. \qquad (17)$$

Equations (16) and (17) can be combined to yield an equation for $D_{local}(\omega_{fi}, r)$,

$$D_{local}(\omega_{fi}, r) = \frac{f_{LSPR}(\omega_l)}{V_{LSPR,l}}\left(1 + \frac{r}{a}\right)^{-2l-4} \mid r > a \;\&\; \theta = 0, \pi. \qquad (18)$$

The density of states for the homogenous medium in the absence of the metal sphere is

$$D(\omega_{fi}) = \frac{\varepsilon_d^{3/2}\omega_{fi}^2}{2\pi^2\varepsilon_0^{3/2}c^3}. \qquad (19)$$

Thus the DOS enhancement is then given by the ratio of Equations (18) and (19), and the total Raman enhancement is the product of this ratio and Equation (15). From these equations we find that both the field and DOS enhancements are maximized when $\omega_l = \omega_0 = \omega_{fi}$. In Raman scattering, these two conditions are approximately met simultaneously due to the fact that the excitation and Stokes emission frequencies do not differ by a large amount. When possible, it is usually optimal to excite SERS using a wavelength slightly blue shifted from the LSPR resonance peak, so that both excitation and emission receive maximum enhancement.

The local DOS effect can be further broken down into the contributions from the dipole LSPR mode, and the quenching effect of the $l \geq 2$ modes. Coupling to the non-radiative modes is more efficient when the scattering molecule is closer to the nanoparticle surface, and may result in a net attenuation of Raman scattering. This results in a trade off wherein molecules that are too close to the metal surface will experience high quenching, while molecules that are far from the surface will not experience a larger local field magnitude. With a carefully designed SERS substrate it is possible to optimize the analyte spacing, the particle size, and the location of maximum field intensity to simultaneously maximize all SERS enhancement contributions.

The final phenomenon briefly considered in this section is the effect of coupled nanoparticles on SERS enhancement. From equation (17) is it evident that as the LSPR mode order increases the effective volume decreases. Furthermore, these high order modes also experience low damping due to lack of radiative coupling. The result is a high D_{local} for these modes, but they cannot be accessed due to their vanishingly small dipole moments. By combining two or more nanoparticles, it is possible to couple the dipole

mode of one particle to the higher order modes of another resulting in high field confinement in the gaps between particles. The field enhancement in the gap between two particles is proportional to Q^2 whereas it is only proportional to Q for an isolated nanoparticle. A qualitative illustration of this effect is provided in Figure 4B. Additionally, the coupling of modes results in a red-shift in the resonant frequency which is advantageous for biomedical Raman application which often use NIR excitation frequencies to minimize absorption and fluorescence.

2.5.1.2Intrinsic and Extrinsic SERS sensor approaches

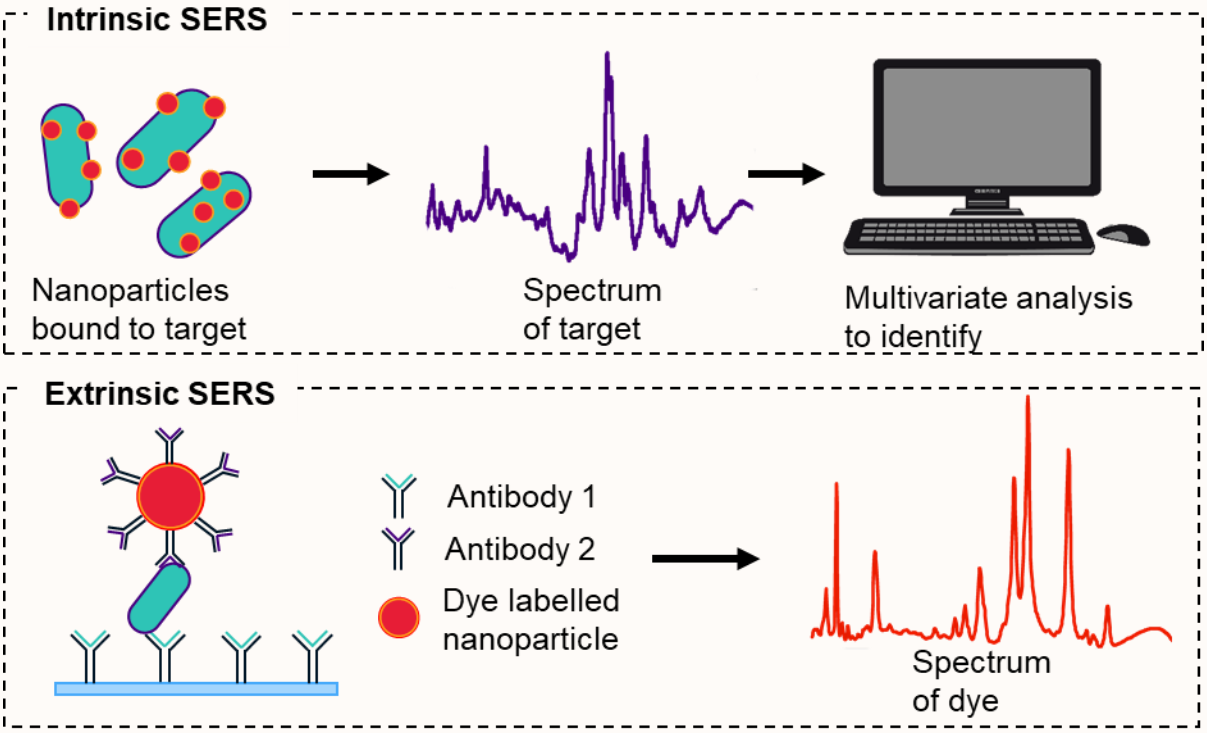

Figure 5: Illustration of intrinsic and extrinsic SERS approaches.

Given a suitable nanomaterial for SERS experiments, there are two general approaches one can take to create a sensor for a specific molecule. The first is the intrinsic approach, which leverages the unique Raman spectral fingerprint of a given molecule, and requires that it bind to the nanostructure within its LSPR field. Upon Raman spectral measurement, the spectrum of the target molecule can be isolated from the spectrum of the background using multi-variate analytical methods. This spectrum can then be used to identify and quantify the molecule. The advantages of this approach are

its simplicity and low cost, as well as the possibility for inherent multiplexing. This is due to the uniqueness of Raman spectra, which facilitates detection of multiple targets at once by deconstructing their superimposed spectra.

Alternatively, one can use the extrinsic approach which relies on measuring the Raman spectrum from a reporter molecule on the surface of a metal nanoparticle which is also bound to a recognition element for the target of interest. This recognition element can be any molecule that binds the target such as antibodies [21], aptamers [22], complementary DNA [23], or molecularly imprinted polymers [24]. The immediate challenge with this approach is the fact that the recognition element is usually much larger than the LSPR field, thus binding the target too far from the metal surface to enhance its intrinsic spectrum. This is the reason the reporter molecule, often a dye with a high Raman cross section, is also applied. One example of this approach is immuno-SERS microscopy, which uses dye-labelled/antibody-conjugated nanoparticles to image the spatial distribution of target antigens in cells and tissue [25]. The major advantage of this approach over intrinsic SERS is the specificity afforded by the recognition element. In the case of immuno-SERS, one will only observe the Raman spectrum of a reporter molecule when the specific target of its associated antibody is present on the cell. The downside to this approach is that the recognition element adds significant cost to the SERS substrate, and multiple elements are required to create a multiplexed detection scheme.

These principals are illustrated in Figure 5 which shows two possible approaches to detect bacteria with SERS. In the intrinsic method, nanoparticles are bound directly to cellular components and the spectra of these are measured. This gives a biochemical picture of the cell which can be used with multi-variate analysis to identify and quantify the bacteria. In the extrinsic scheme, bacteria are first captured by a primary antibody bound to some substrate. Subsequently, a secondary antibody bound to a dye-labelled nanoparticle will also bind to the bacteria. Therefore, one measures the spectrum of the reporter dye, and its intensity reflects the quantity of the bacterium targeted by the antibodies. Biorecognition can also lead to analyte enrichment in some detection schemes, thereby increasing the measured signal and lowering the detection limit. For example, consider the SERS lateral flow assay to detect HIV DNA developed by Fu *et al*

[26]. Lateral flow assays are a common type of disposable biosensor built on a highly porous substrate such as nitrocellulose. Well known examples of which are over the counter pregnancy test strips, or rapid COVID-19 antigen tests [27]. These sensors consist of three main regions; the conjugate pad containing the bio-conjugate SERS substrate, the test line containing a primary antibody, and a control line containing a secondary antibody. As the sample flows through the porous membrane via capillary action, the target of interest will be concentrated in the test line due to capture by the primary biorecognition element.

The work in this thesis will primarily utilize intrinsic SERS to create medical diagnostic systems in order to take advantage of inherent multiplexing by intrinsic spectral differences. Machine learning algorithms will be leveraged in order to extract these differences and develop regression/discriminant models. These algorithms are the subject of Chapter 3.

2.5.1.3 Fabrication or synthesis of SERS substrates

When developing SERS based methods for chemical detection/quantification it is important to consider the fabrication route used. Ideally, a SERS substrate has high affinity to the target analyte, a strong plasmon resonance band at the excitation frequency, high stability and long shelf life, and a facile, inexpensive, reproducible fabrication method. Conventional approaches to generating SERS substrates fall into one of several categories. Broadly speaking, metal nanostructures can be created through either top-down or bottom-up approaches [28]. In the bottom-up approach, one often starts with a solution of metal salts such as Au^{3+} or Ag^{+} which are then reduced to atomic metals and formed into nanostructures in the presence of templates formed by surfactant molecules. The top-down approach starts with macroscopic metals which are either chemically etched, or subject to physical lithography, to form nanostructures. It is also possible to combine these approaches in order to fabricate substrates with novel properties.

The chosen fabrication method will result in either a substrate bound static structure, or a more dynamic colloidal suspension of particles. Substrate bound nanostructures tend

to have better reproducibility and stability than their colloidal counterparts, and are frequently formed through top-down shaping of a supporting substrate followed by metal deposition. Examples of this include silicon nanopillar arrays formed by reactive ion etching followed by electron-beam evaporation of gold [29], and magnetron sputtered silver on glass substrates with nanoscale features from laser ablation [30]. Colloidal approaches tend to be much simpler, less expensive, and possess a higher surface area with which to interact with analytes in solution. These colloids are usually formed bottom-up through a chemical synthesis, and may aggregate around target molecules to create strong coupled plasmonic resonances. This chemical synthesized colloid approach is the one taken throughout this thesis.

There are four main ingredients to synthesize metal nanoparticles; the metal ions, the reducing agent, the capping agent, and the surfactant [31]. The reducing agent is the molecule which donates electrons to metal ions in solution in order to convert them to their atomic form. The capping agent is the molecule which binds to the surface of the nanoparticles to stabilize the colloid, either by imparting surface charge on the particle thereby electrostatically repelling nearby particles, or by imparting hydrophilicity then providing each particle with a protective layer of water or other solvent molecules. The surfactant molecule essentially forms a nanoscale template through interactions with the metal surface, thereby guiding the shape in which the nanostructure grows. In the most common synthesis approach for gold nanoparticles, known as the citrate or Turkevitch method, citrate acts as the reducing, capping, and surfactant molecule when it is added to a boiling solution of metal ions [32]. In the reaction, citrate is oxidized to dicarboxy-acetone, acetone, and other related molecules to reduce the gold ions, and also convert them to a less reactive hydroxylated form. The initially reduced gold atoms form small clusters and eventually seed particles. The remaining metal ions are attracted to the electrical double layer of the seeds where they diffuse, concentrate, and are reduced onto the surface thereby growing the particles [32]. The end result is a monodisperse colloid of approximately spherical particles each with a layer of adsorbed citrate which imparts a negative surface charge. In a similar method, silver nanoparticles can also be reduced by citrate, but there are some critical differences between these two routes. From a practical level, all solutions involving colloidal silver should be deoxygenated because silver is

highly susceptible to oxidation. This reduces conductivity and thereby attenuates the surface plasmon resonance. Furthermore, unlike citrate reduction of gold, production of citrate-silver nanoparticles results in a highly polydisperse solution of nanoaggregates [32].

Citrate is only weakly adsorbed to the surface of nanoparticles after synthesis, and it is therefore possible to replace it with a different capping agent. Changing the capping agent changes the physiochemical properties of the nanoparticle, allowing it to bind/repel different molecules, providing additional stability, and potentially shifting the plasmon resonance band slightly. A good way to displace adsorbed citrate is to introduce a molecule with a high affinity for the metal surface, such as molecules with thiol (SH) or amine (NH_2) moieties. An example of a commonly used capping agent is thio-(polyethylene glycol) (PEG-SH), which has a high affinity for gold surfaces due to the favourable formation of Au-S bonds [33]. PEG-SH makes the nanoparticle surface hydrophilic, providing excellent colloidal stability in water. This also has the result of significantly changing the way in which proteins in biological solutions interact with the particle. When nanoparticles are introduced into a biological fluid, a shell of protein known as a “protein corona” quickly forms around them [34]. The hydrophilic surface generated by PEG-SH capped particles causes preferential binding of water molecules and thus less protein adsorbs to the surface. This is one example, but one should expect that any change in the capping agent will result in different interactions in complex biological media. Modulation of nanoparticle surface chemistry through biological interactions is applied in Chapter 5 in developing a chemoresistance sensor for ovarian cancer. Tailoring the surface chemistry of particles for better performance in other applications is discussed more in Chapter 6.

It is also possible to generate different shaped nanoparticles by adding different surfactants or etchants into the synthesis. For example, it is possible to make silver nano-plates with a modified citrate synthesis by adding hydrogen peroxide as an etchant [35]. Many shapes can be made such as cubes [36], rods [37], stars [38], dendrites [39] and more. The purpose of using these more complex shapes is the nature of the plasmon resonance. Sharper features, such as the corners of a nanoplate or the tips of a star tend

to have strong resonances at longer wavelengths with higher Purcell factors than the dipole mode of a spherical particle [40–42].

2.5.2 Fiber Enhanced Raman Scattering

Conventionally, Raman spectroscopy is performed using a focused Gaussian beam with the sample contained inside a glass or quartz cuvette. Since Raman scattering efficiency is proportional to the incident intensity, the majority of the measured spectrum will originate from the focal point. The volume of this region is twice the beam's Rayleigh range long, with a radius approximately equal to the beam waist. However, due to the nature of a Gaussian beam, there is a trade-off between the interactions length and the cross sectional area of the beam. For example, if one wishes to increase the interaction length then the beam waist must also grow thereby reducing the maximum intensity at the focus [43]. A figure of merit has been defined in order to illustrate this tradeoff between interaction length and focal spot size.

$$F_{om} = \frac{L\lambda}{A_{eff}} \qquad (20)$$

This value is always equal to 2 for a Gaussian beam. In order to achieve higher figures of merit, we can employ a hollow optical fiber which exhibits low losses and a small core diameter [44]. Previous research by our group demonstrated the efficacy of a hollow-core photonic crystal fiber (HC-PCF) as a means of increasing Raman scattering by co-coupling light and liquid sample into the same cavity [45]. Photonic crystal fibers exhibit a periodic array of air holes around a central defect which creates an incomplete bandgap capable of guiding light within the core. These fibers exhibit extremely low losses (<0.03 dB/m for HC1550 from Thorlabs), which means that the length L in equation (20) is effectively equal to the fiber length. Furthermore, A_{eff} is constant over the length of fiber and is simply equal to the core area. Thus it is now possible to increase the figure of merit proportional to the length of an HC-PCF.

For HC-PCF Raman enhancement, the fiber can be filled with the desired analyte either by selectively filling only the core, or by non-selectively filling the core and photonic crystal. In the latter case, filling the crystal with a material of higher refractive index results

in a blue-shift of the bandgap, thereby allowing the fiber to guide shorter wavelengths [46]. The new peak wavelength in the band gap is given by

$$\lambda = \lambda_0 \left[\frac{1 - \frac{n_h^2}{n_c^2}}{1 - \frac{n_{h0}^2}{n_c^2}} \right]^{1/2} , \qquad (21)$$

where n_h and n_c are the refractive indices of the holey region and cladding respectively, with n_{h0} being the design index with guided wavelength λ_0. For example, the HC-PCF used in this thesis normally guides 1550 nm when filled with air and has a cladding made of silica (n_c = 1.45). When the fiber is filled with water (n_h = 1.34), then from equation (23) the new guiding wavelength is ~818 nm. The band gap is quite broad, and spans the full fingerprint range of Raman scattered frequencies when excited at 785 nm. The primary advantage of non-specific filling is ease of implementation. The HC-PCF can simply be cleaved to the desired size and inserted into the filling mechanism. However, it is necessary to consider that the photonic crystal is very fragile and can be damaged by solid material flowing through it. Furthermore, the capillary pressure in the crystal lattice is very high due to their small diameter, meaning that any air bubbles trapped inside are extremely difficult to remove. These will disrupt the photonic bandgap and possibly eliminate any guiding.

Alternatively, it is also possible to selectively fill the core, thereby improving the robustness of the fiber at the cost of additional complexity. Waveguiding is now a result of total internal reflection rather than a photonic bandgap due to the low effective index of the crystal. There are several methods of selectively filling the core, including the use of adhesives, back-pressure, or collapsing the crystal at the cleaved ends of the fiber [47–49]. This last method was recently developed by Yan *et al* and is probably the most practical of these approaches [50]. In this process, a section of the photonic crystal is collapsed using the plasma from a conventional fiber splicing machine. Then the fiber is cleaved in the tapered region where the air holes of the crystal are collapsed while the central defect is still open. This cleaving stage involves the use of a cleaver with a high quality blade and integrated imaging optics so that the fiber can be cut at the correct

position. In this thesis, all experiments have employed the non-selective filling approach for the sake of ease of implementation. In Chapter 4, the possibility of fiber modification to improve the performance of the system is discussed, including selective filling of the core.

2.5.3 Other Enhancement Methods

There are several other methods to enhance Raman scattering that are not employed in this thesis, but some popular methods will be mentioned here for completeness. These include resonant and non-linear enhancement to Raman scattering, and controlling the chemical enhancement component of SERS using electrodes.

2.5.3.1 Resonant Raman scattering

Another method to enhance Raman scattering is resonant Raman spectroscopy, wherein scattering is excited by a wavelength whose energy lies near an electronic transition in the molecule of interest. Consider equation (7), the denominator of which contains the term $\omega_0 - \omega_{ei}$, implying that the Raman scattered rate will be very high if $\omega_0 = \omega_{ei}$. This is only the case if the transition from $|i\rangle$ to $|e\rangle$ is an allowed electronic transition [3]. This may cause both resonant Raman scattering and fluorescence to occur simultaneously, albeit at different time scales and frequencies. Due to the fact that real electronic transitions are required, resonant Raman is conventionally only used when assaying chromophores. For example, this approach has been used to assay hemoglobin in red blood cells because the heme molecules in this protein strongly absorb green light [51].

2.5.3.2Electrochemical SERS

One can also take a substrate-supported SERS platform, and integrate it into an electrochemical cell resulting in electrochemical SERS (EC-SERS). Using this electrode, one can cause additional chemical SERS enhancement, cause EC-SERS measurable redox reactions, and attract or repel molecules of a certain charge from the surface. Furthermore, SERS substrates can be grown via electrodeposition of a variety of metals allowing for the *in situ* fabrication of SERS active sites. Therefore, these systems gain many of the benefits of both electrochemical and SERS sensors. There are several

examples of this approach in the literature such as the work by Karaballi *et al.* who developed an EC-SERS system to detect DNA hybridization events, and they applied this to the diagnosis of tuberculosis (TB) since this disease is associated with a DNA fragment biomarker that is expelled in the urine [52]. The substrate used was a screen-printed carbon electrode functionalized with gold nanoparticles onto which target complimentary DNA was bound. The key to this sensor is the fact that the TB DNA contains adenine residues, whereas the capture probe on the electrode does not. Due to electrochemical effects, the SERS signal of adenine is amplified when the surface voltage is set to -0.8 V. This is an example of EC-SERS being used to modulate the electrochemical SERS enhancement effects.

2.5.3.3 Non-linear Raman

There are also non-linear optical approaches that can be used to enhance Raman scattering such as stimulated Raman scattering (SRS), or coherent anti-Stokes Raman (CARS). These effects arise from the coherent excitation of a vibration when two excitation wavelengths, ω_l and ω_s, are incident on a sample whose frequencies differ by the vibrational frequency of a certain bond mode (ω_v). The mathematical descriptions of these processes are intuitively derived in Dr. Robert Boyd's book on non-linear optics, an extremely brief summary of which is presented here [53]. For SRS, the non-linear susceptibility of the Stokes scattering polarization (χ_R) is given by

$$\chi_R(\omega_l, \omega_s, \omega_v) = \frac{\varepsilon_0(N/6m)(\partial\alpha/\partial q_k)|_0}{\omega_v^2 - (\omega_l - \omega_s)^2 + 2j(\omega_l - \omega_s)\Gamma},$$

which is purely imaginary and negative when $\omega_v = \omega_l - \omega_s$, thereby exhibiting optical gain. There is another component of the total non-linear polarization that results from a four wave mixing process at the anti-Stokes frequency for the vibration of interest (ω_{as}) which we define as $\chi^{(3)}(\omega_{as} = 2\omega_l - \omega_s) \equiv \chi_F(\omega_{as})$. This is the resonant part of the CARS susceptibility given by

$$\chi_F(\omega_{as}) = \frac{\varepsilon_0(N/3m)(\partial\alpha/\partial q_k)|_0}{\omega_v^2 - (\omega_l - \omega_s)^2 + 2j(\omega_l - \omega_s)\Gamma},$$

which clearly has the same resonance condition as SRS. However, under most conditions SRS will dominate over CARS because the former is automatically phase matched whereas we require the wave-vectors in CARS to satisfy $k_{as} - (2k_l - k_s) \rightarrow 0$. Under perfect phase matching, there is coupling between the Stokes and anti-Stokes modes resulting in small SRS and large CARS. As this brief introduction shows, both SRS and CARS non-linearly enhance Raman scattering from a single vibrational mode. This means that a tunable source for ω_l or ω_s is required in order to generate a full spectrum.

Chapter 3 Raman Chemometrics with Support Vectors Machines

3.1 Introduction

Raman spectra represent unique molecular fingerprints, thus they can be effectively utilized in two very important tasks; discrimination between different chemical samples, and quantification of target molecules. In order to successfully perform either, a robust data analysis method is required. Biomedical Raman data have additional complexity due to non-linearities in biological spectra. These may arise from autofluorescence [54], or variation of scattering and absorption coefficients in tissue and cellular samples [55]. To overcome these challenges, several multivariate analysis (MVA) techniques have been applied to classify or quantify samples based on their Raman spectrum. Multivariate tools are required because spectral data has a large number of variables, or dimensions, associated with them. This will cause an ordinary least squares solution to fail, or produce an over-fitted model. One common technique, projection to latent structures or partial least-squares (PLS), is commonly applied to Raman data. This technique reduces the dimensionality of the input data by projecting them onto axes, known as latent structures, which describe the highest degree of covariance between the input and output data [56]. It has been demonstrated that PLS produces excellent regression and discriminant models [57]. For example, Bratchenko *et al* demonstrated that Raman spectra of skin could be predictive of kidney failure using PLS discriminant analysis [58].

However, PLS may fail to produce a good model when a non-linear solution is required. One method to solve for these functions is to first project the data into a higher order feature space where conventional linear methods can then be applied. Take for example an algorithm trying to solve for a circular discriminant surface. In this case there is no line that can satisfactorily solve the problem. However, if the data were first to be projected from two dimensions to three by applying a quadratic transformation, then the solution in this new feature space is a linear plane. If this solution is then projected back into the original data space, then it takes the form of a circle as expected. Instead of

explicitly calculating these new feature vectors, one can apply the so called “kernel trick”, to find a solution using only a function of the original data. One of the more popular techniques to utilize kernel functions is the Support Vector Machine (SVM), developed by Vladimir Vapnik [59]. This technique has already been applied in the analysis of biomedical Raman spectra. For example, SVMs in conjunction with Raman spectroscopy have been used to diagnose hepatitis B infections from blood plasma [60], genotype brain tumors [61], and diagnose Alzheimer’s disease from cerebrospinal fluid [62]. However, SVMs possess a number of “hyperparameters” that cannot be trivially selected in order to find a good model. These parameters are often selected by a knowledgeable user, or found in a brute-force full factorial method known as grid search [63]. There are a number of numerical optimization techniques that may be applied in place of grid search, such as genetic algorithms, or particle swarm optimization. This chapter outlines the deployment of complex kernel functions with many hyperparameters, optimized by a genetic algorithm (GA). This GA uses an analytical fitness function which can be computed quickly, and is minimized when an SVM model is both accurate and robust. The resultant SVM models are optimized much faster than grid search or GAs which use n-fold cross validation as a fitness function. These non-linear models outperform conventional PLS analysis in developing regression models for biomedical Raman spectra, and in Chapter 4 we will show that SVMs outperform PLS in discrimination analysis as well.

This chapter begins with the theory behind SVMs and GAs in Section 3.2. Section 3.5 contains the validation of our GA-SVM algorithm on standard test data. In Section 3.6, the GA-SVM will be applied to a non-linear regression problem involving Raman spectra. These were made by adding polystyrene microspheres to different concentrations of ethanol. Finally in Section 3.8, GA-SVM was compared to PLS using regression and classification problems built from our lab’s database of biomedical Raman spectra. Specifically, this included different concentrations of heparin [64], γ-aminobutyric acid (GABA), or glutamate in serum [65], and different cell densities of leukaemia cells [66]. To our knowledge, this was the first time that an analytical fitness function was used in GA-SVM regression, and certainly the first time for biomedical Raman datasets. This chapter will conclude with possible research directions to improve upon our machine learning algorithms for Raman data processing.

3.2 Regression, linear discriminants, and PLS

Before exploring a new approach to Raman chemometrics, it is valuable to first understand the classical approaches. As discussed previously, the intensities of Raman bands associated with a certain molecule are proportional to the number of molecules in the excitation field, and different molecules possess unique sets of vibrational frequencies. Thus the most obvious approach to quantification or differentiation using Raman spectra is to select only the bands associated with the molecule of interest to develop a model. Consider the ethanol spectrum shown in Figure 2 in the previous chapter, the strongest Raman band is the C-C-O symmetric stretch resonance at 880 cm^{-1}. Using only this band intensity, it is possible to quantify dilutions of ethanol in water very easily using a normal least squares regression approach. If the measured intensities are concatenated into a vector $\vec{x} = [x_1 \quad \cdots \quad x_n]$ with associated ethanol concentrations $\vec{y} = [y_1 \quad \cdots \quad y_n]$, then these are simply related by a linear model $\vec{y} = \vec{x}w + b$. One can also include multiple Raman bands in the regression (for example 880, 1049, 1100, and 1446 cm^{-1} for ethanol) by storing the intensities in a matrix X wherein each row contains a different sample and the columns are the measured values for each. Thus if there are n samples and m Raman shifts per sample, then X is an $n \times m$ matrix. The goal is to find a slope vector $\vec{w} \in \mathbb{R}^m$ which minimizes the squared error of predicting the output $\vec{y}$ which is given by

$$\hat{\Re}(y, w, x) = \sum_{i=1}^{n} (y_i - x_i w)^2. \qquad (22)$$

Assuming that this is a well-conditioned convex optimization problem, the minimum is a unique solution where the first derivative of the squared error with respect to each element in w is zero ($\partial\hat{\Re}/\partial w_j = 0$). The resultant normal equations can be written in matrix notation as

$$X^T X w = X^T y$$
$$w = (X^T X)^{-1} X^T y. \qquad (23)$$

The final result is a matrix inversion problem to solve for w, which is well conditioned when the number of samples exceeds the dimensionality of the data ($n > m$). However,

if $m > n$ then the problem is ill-conditioned and there may be zero or infinite possible solutions. This is the issue encountered frequently when analysing complex Raman data such as biomedical spectra, where the whole spectrum is often used. This is often the case because it may be unknown in advance where the Raman bands of interest lie, or in complex samples the bands of interest could span the entire spectral range.

PLS solves this problem of dimensionality by projecting data onto a new set of axes in $\mathbb{R}^{p<m}$ where ordinary least squares can then be performed. The value p is usually known as the number of PLS components. This new set of axes, known as "latent structures", are the directions in the original feature space which describe the highest degree of covariance between the input and output matrices X and Y. If T $(n \times p)$, known as the "X scores", is X projected onto these latent variables then an ordinary least squares regression in terms of T is simply

$$Y = TC^T + F. \quad (24)$$

Where C is the regression matrix and F are the residuals. To generate these scores, mean centred X and Y are decomposed through

$$X = TP^T + E \quad (25)$$

$$Y = UQ^T + F. \quad (26)$$

Where U are the Y scores, E are the X residuals, and P $(m \times p)$ and Q $(k \times p)$ are the X and Y loadings respectively. In essence, the loading vectors are the projection matrix from the mean centred data onto the latent structures such that $XP \cong T$ and $YQ \cong U$. Each individual loading $\vec{t}$ and $\vec{u}$ are found such that

$$\left[cov\left(\vec{t}, \vec{u}\right)\right]^2 = \max\{[cov(X\vec{r}, Y\vec{s})]^2\}|_{|r|=|s|=1}, \quad (27)$$

which is performed iteratively through a solving algorithm. MATLAB specifically uses the SIMPLS algorithm by Sijmen de Jong [67]. The end result in terms of the original data is

$$Y = X\beta + F^*, \beta = X^T U(T^T X X^T U)^{-1} T^T Y. \quad (28)$$

The matrix $T^T X X^T U$ is a non-singular square matrix with dimension $p \times p$, and thus it has a unique inverse and the regression matrix β can always be found. It is possible to also use PLS for discriminant analysis (PLS-DA) for g classes by encoding a class in Y as $y_i = [z_1 \quad \cdots \quad z_g]$ such that $z_j = 1$ if y_i is class j and zero otherwise. Dimensional reduction in PLS-DA is closely related to Fisher linear discriminant analysis [57].

It is also worth noting, that $\vec{t}$ and $\vec{u}$ are solved for in order of decreasing explained covariance from the first to the p^{th}. There are two important results from this. First is that there is a certain value of p for which solving for any more latent structures would be superfluous because effectively all of the covariance has been explained. Finding the optimal number of components is usually done by iteratively increasing p and calculating the error of the resultant PLS model. The error will initially decrease, reach a minimum, and then increase, thus the number of components associated with the minimum is selected as the optimum. The error begins increasing as higher orders of p begin to describe noise in the data, which is detrimental to the PLS model. Secondly, by analysing the loading vectors it is possible to estimate which variables contribute the most towards the covariance between X and Y. Particularly the first loading, which described the highest degree of covariance, can yield valuable information about the regression or discriminant problem at hand. The magnitude of the loading at a certain Raman shift indicates its importance in the PLS model.

3.3 Theory of Support Vectors Machines and Genetic Algorithms

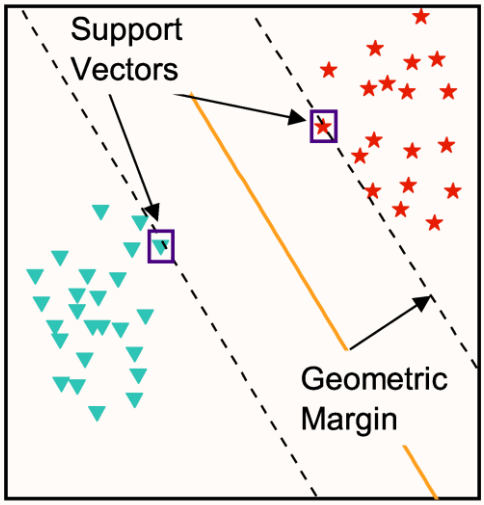

Figure 6: A simple binary classification problem wherein one seeks a hyperplane that optimally separates the two classes.

3.3.1 Classification by SVM

Consider a simple binary classification problem like the one shown in Figure 6. We seek a decision boundary that separates class 1 (the triangles) from class 2 (the stars), but this problem is poorly defined. It is evident that there are infinite possible lines that separate these two classes. However, there exists an optimal solution wherein the distance from the discriminant surface to the closest data points, the geometric margin, is maximized. SVMs are based on finding this optimal solution, which is entirely described by the points on the geometric margin, also known as support vectors [68]. Points distant from the geometric margin are not needed, hence SVMs generate a sparse solution to the discrimination problem. Empirical risk ($\widehat{\Re}$), is minimized in order to find the optimal discriminant hyperplane. In our current example, let all data points be contained in a matrix $X = [x_1\ x_2\ \dots\ x_n]^T$ consisting of n samples $x_i \in \mathbb{R}^d$. For each point, a binary class label $y_i \in \{-1,1\}$ is then assigned. For example, let the triangular points have $y_i = 1$ and the stars $y_i = -1$. In SVMs $\widehat{\Re}$ is given by the hinge loss function

$$\widehat{\Re}(f(x), y) = \sum_{i=1}^{n} \max[0, 1 - y_i f(x_i)], \tag{29}$$

where $f(x_i)$ is the SVM output for point x_i. If the prediction is correct, then the hinge loss will be zero because y_i and $f(x_i)$ will have the same sign, thus causing $1 - y_i f(x_i)$ to be negative. If the prediction is incorrect then y_i and $f(x_i)$ will have opposite signs, and the hinge loss will be some positive number. Let the possible linear solution to this problem be of the form $f(x_i) = \langle x_i, w \rangle + b$, where w is the hyperplane weight vector and b is its offset. Additionally, add the norm of the hyperplane ($\frac{1}{2}\|w\|^2$), to regularize the problem which will allow for solutions to ill-conditioned problems [69]. This is necessary in most Raman spectroscopy applications, because the dimensionality of the data usually far exceeds the number of samples. A min-norm hyperplane decreases the complexity of the solution to avoid overfitting and improve generalizability. The new minimization problem is formulated as

$$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \max[0, 1 - y_i(\langle x_i, w \rangle + b)], \tag{30}$$

wherein C, known as the "cost" or the "box constraint", is the hyperparameter which determines the trade-off between errors and a small norm hyperplane [59]. Analytically, there is no way to find the minimum for this function, because the $\max[]$ operator is not differentiable. To avoid this problem, some error is allowed for by introducing slackness ξ_i which is the error of prediction on point (x_i, y_i). The minimization problem is now

$$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \xi_i$$
$$\text{such that } \begin{cases} y_i(\langle x_i, w \rangle + b) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases}. \tag{31}$$

Allowing for some error results in a "soft-margin" classifier, which allows for convergence in problems that are not perfectly separable. This optimization is easier to solve in the dual form of its Lagrangian [68], thus we now seek to minimize

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j\langle x_i, x_j\rangle - \sum_{i=1}^{n}\alpha_i$$

$$\text{such that}\begin{cases} 0 \le \alpha_i \le C \\ \sum_{i=1}^{n}\alpha_i y_i = 0\,, \end{cases} \qquad (32)$$

where α_i is the i^{th} Lagrange multiplier. SVMs solve for these Lagrange multipliers instead of w directly, which is related to the Lagrange multipliers by $w = \sum_{i=1}^{n}\alpha_i y_i x_i$.

The magnitude of a Lagrange multiplier α_i can be thought of as the importance of x_i in determining the discriminant surface. Therefore, all support vectors which by necessity contribute to the final solution will have $\alpha_i > 0$. For a certain support vector x_i the associated α_i is high if the point is on or within the soft geometric margin, or is erroneously classified. Thus, if the sum of the α_i's is small, then the discriminant has a minimum number of poorly classified support vectors. Also note that C sets an upper limit on the Lagrange multipliers, thereby allowing for some errors in order to improve robustness.

3.3.2 Multi-class models

To extend SVM classification from binary to multi-class cases there are two common methods, "one versus one" (OVO) and "one versus all" (OVA). Both of these approaches construct multi-classifiers out of a set of binary models. As the names imply OVO uses binary classifiers covering all unique combinations of classes, while OVA uses binary classifiers that compares one class against all others at once. For example, consider a classification problem used to differentiate between three different species of bacteria (as in Chapter 4). For now, name these species *A*, *B* and *C*. In OVO there would be three binary classifiers *A* vs. *B*, *A* vs. *C*, and *B* vs. *C*. Similarly, in OVA there are also three binary classifiers *A* vs. *B* and *C, B* vs. *A* and *C*, and finally *C* vs. *A* and *B*. A new sample is classified by applying each of these three models, and the final class is determined by the output with the largest magnitude.

3.3.3 Regression by SVM

Linear regression is very similar to the linear classifier case except that a different risk function is used, namely Vapnik's ε-insensitive loss function [70], which is given by

$$\widehat{\Re}(f(x), y) = \sum_{i=1}^{n} \max[0, \left|\left(y_i - f(x_i)\right) - \varepsilon\right|]. \quad (33)$$

As the name suggests, errors smaller than ε are ignored by this function, allowing for some slackness as per SVM classification. Define ξ and ξ^* as errors greater than ε above and below the regression line respectively. The SVR algorithm thus attempts to minimize

$$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i{}^*),$$

$$\text{such that} \begin{cases} y_i + \langle x_i, w\rangle - b \le \varepsilon + \xi_i \\ \langle x_i, w\rangle + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i^{(*)} \ge 0 \end{cases}, \quad (34)$$

where the cost C possesses the same function as before. Again, this constrained optimization is easier to solve in the dual form Lagrangian which is given by [70]

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha_i - \alpha_i^*)\left(\alpha_j - \alpha_j^*\right)\langle x_i, x_j\rangle + \varepsilon\sum_{i=1}^{n}(\alpha_i - \alpha_i^*) - \sum_{i=1}^{n} y_i(\alpha_i - \alpha_i^*),$$

$$\text{such that} \begin{cases} \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0 \\ 0 \le \alpha_i^{(*)} \le C \end{cases} \quad (35)$$

In the regression case, there are now two Lagrange multipliers, α_i and α_i^* for each support vector x_i. The weight vector for the linear regression hyperplane is constructed from a linear combination of the support vectors as $w = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)x_i$ and the final solution is

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*)\langle x_i, x\rangle + b. \qquad (36)$$

Non-support vectors ($\alpha_i^{(*)} = 0$) are those points which exist within the ε-tube near the regression line. Again we arrive at a sparse solution as these points can be removed. Support vectors are those points whose regression error is greater than ε, thus we arrive at the same conclusion as in the classification case; fewer non-zero Lagrange multipliers is related to an accurate solution because more points have an error less than ε.

3.3.4 Non-linear Solutions

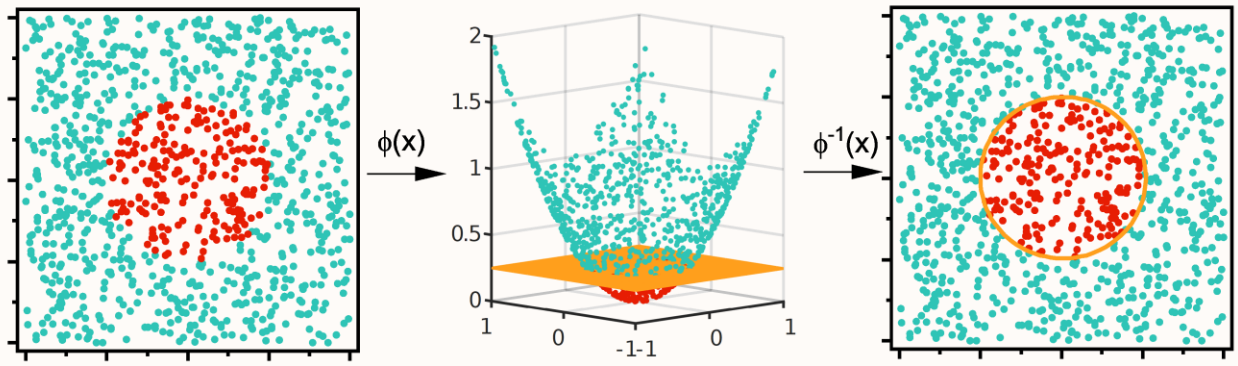

Figure 7: Example of projecting data into a higher order feature space where one can find a linear solution.

In order to learn non-linear models using SVM, we may project the data into a new feature space where the solution becomes linear and then solve the Lagrangian as before. Let $\phi(x_i)$ be a mapping function from $\mathbb{R}^d$ to $\mathbb{R}^k$ such that $k > d$. Now x_i can be replaced with $\phi(x_i)$ in the optimization problems from the preceding sections to build a linear model in a high dimensional feature space. This is then equivalent to a non-linear solution in the original data space. An example of this process is shown in Figure 7, where we clearly seek a circular discriminant surface to differentiate the two colors of dots. In order to solve this problem with a linear solution, it is first projected into a quadratic feature space. Now, the classes can be separated by a hyperplane which becomes a circle when projected back into the original data space. Unfortunately, Raman data already consists of hundreds of features, meaning that computing the new feature space would be computationally expensive or even impossible. To avoid this problem, a method known

as the “kernel trick” can be used. Both of the Lagrangians derived earlier rely on inner products within x, which can be replaced with $\langle x_i, x_j \rangle = \langle \phi(x_i), \phi(x_j) \rangle$ after mapping. For a certain feature space, inner products can be computed by a kernel function $\kappa(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ which uses the original data as inputs. Thus, by using this kernel function in place of the inner products in Equations (32) and (35), explicit computation of $\phi(x_i)$ is avoided. There are many possible kernel functions that may be used to represent different feature spaces. Three basic functions are employed throughout this thesis, the polynomial kernel κ_{poly}, the radial basis function (RBF) kernel κ_{RBF}, and the rational quadratic kernel κ_{RQK}. These are given by

$$\kappa_{poly}(x_i, x_j) = (\gamma \langle x_i, x_j \rangle + c_0)^p \quad (37)$$

$$\kappa_{RBF}(x_i, x_j) = e^{-2\sigma^{-2}\|x_i - x_j\|^2}. \quad (38)$$

$$\kappa_{RQK} = \sigma^2 \left(1 + \frac{(x_i - x_j)^2}{2\alpha l^2}\right)^{-\alpha}. \quad (39)$$

In addition to the input data x, there are a number of hyperparameters required by each kernel function. For κ_{poly} there is the scaling term γ, the offset c_0, and the polynomial order p. For κ_{RBF}, which represents a Gaussian feature space, the only parameter is the variance σ^2. Finally, κ_{RQK} represents an infinite sum of exponentiated quadratic kernels with variance σ^2, length scale l, and scale variance α. Kernels may also be combined by multiplication or addition, which may be interpreted as AND/OR logic respectively. When the underlying patterns in the data are unknown, optimizing sums of kernels allows the user to cover many possible non-linear interactions. For example, in this chapter we will employ a combined kernel κ_{combi} given as a linear combination of the polynomial and RBF kernels

$$\kappa_{combi}(x_i, x_j) = A_1(\gamma \langle x_i, x_j \rangle + c_0)^p + A_2 e^{-2\sigma^{-2}\|x_i - x_j\|^2}. \quad (40)$$

3.3.5 Hyperparameter Optimization

All SVMs have at least one hyperparameter, the box constraint C, in addition to all of the parameters of the chosen kernel. In the absence of an expert user who is intimately familiar with both the data and SVMs, and is therefore able to manually select appropriate parameters, we must instead employ an optimization method. The most basic approach is to perform a brute force, full-factorial, optimization by trying all possible combinations of hyperparameters. This technique, known as grid-search, then compares each model on the basis of the root mean squared error of prediction (RMSEP) for regression models

$$\mathrm{RMSEP} = \frac{\sum_{\mathrm{i}=0}^{\mathrm{N}} (\mathrm{y_{t,i}} - \mathrm{y_{p,i}})^2}{\mathrm{N}}, \tag{41}$$

where $\mathrm{y_{t,i}}$ and $\mathrm{y_{p,i}}$ are the true and predicted outputs for x_i respectively and N is the number of samples. For classification problems, models are compared on the basis of accuracy

$$Accuracy = \frac{TP + TN}{N}, \tag{42}$$

where TP and TN are the true positive and true negative rate respectively. Furthermore, when training any machine learning model, the total dataset must be divided into training and validation subsets. As the name implies the validation set is used to estimate the final performance of the model, and must therefore not be used in any of the preceding optimization. To calculate accuracy or RMSEP using only the training set, n-fold cross validation is used. In this technique, the training set is divided into n subsets (also known as folds) and the model is trained on all but one, and the remaining fold is used to validate the model. The final performance estimate is then given as the average error/accuracy after each of the n-folds has been excluded from training. As a result of this, the grid search algorithm must solve the SVM optimization problem nN^m times, wherein N is the number of hyperparameters and m is the number of values tested per parameter. For smaller kernels such as the RBF function, which only has two hyperparameters, this may not be overly computationally expensive. However, if we consider 10-fold cross validation for κ_{combi} which has seven hyperparameters, the optimization time explodes for higher values of m, and this quickly becomes unreasonable. Using a very coarse grid (low values

of m) to save time, may result in missing the optimum point. As a result, using complex kernels with grid search is undesirable and should be replaced with a less intensive search algorithm.

3.3.6 Genetic Algorithm and Fitness Functions

Instead of using grid search to optimize hyperparameters, this thesis employs a classical genetic algorithm which is modelled off a simplified version of natural selection. In this technique, hyperparameters are coded in a "genome" which in this case is a concatenated string of binary values for each parameter. An initial "population" of possible models is generated using randomly generated sets of hyperparameters, and these are compared on the basis of some fitness function. Population members with better fitness are able to "reproduce" with one another wherein the genome of two parents is split and then crossed to make two new children. There is also the possibility for random "mutations" to occur, where a binary digit is inverted. This allows populations to explore the entire solution space, and avoid falling into local optima. As population members with poor fitness are removed, the overall fitness of the population increases, thus converging to some optimum.

The fitness function represents the desired output of the optimization. In the case of finding good SVM models, this is high prediction accuracy. One possible fitness function is to use the n-fold cross validation output as described in the previous section. However, this can still be computationally expensive since the Lagrangian must be solved n times per population member. Instead, Runarsson and Sigurdsson have proposed one fitness function for a genetic SVM classifier, which takes advantage of the rigorous mathematical basis for these learning machines (in contrast to models without an analytical basis such as neural networks) [71]. This fitness function is given by given by

$$fitness_{classification} = \left(R^2 + {}^1/_C\right)\sum_{i=1}^{n}\alpha_i \ . \qquad (43)$$

This function is low when the classification model is both robust and accurate, thus the resultant genetic algorithm attempts to find a model associated with a minimum fitness

[71]. The rationalization for this function comes from the underlying theory discussed in Section 3.3.1. The sum of multipliers is large when there are many support vectors (over fitting), and when the support vectors are erroneously classified. Therefore, this term in the fitness function represents both the robustness and accuracy of the SVM. The R^2 term is the radius of the hypersphere containing all of the data points about the origin in the high dimensional feature space and can be approximated by

$$R = \max_{1 \le i \le n} [K(x_i, x_i)], \tag{44}$$

where $K(x_i, x_i)$ is a matrix whose values are the kernel products $\kappa(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ [71]. For a robust solution, the sub-space containing all of the features of the data should be small, lest the machine over-fit the data by using large subspaces of the features. Finally, the $1/C$ term is included to penalize small box constrains. Recall that no α_i can be greater than C, thus $\sum_{i=1}^{n} \alpha_i$ may be artificially minimized with a small C simply by restricting the size of the α_i's. To extend the use of this fitness function to SVM regression, we can apply the same logic and simply add the other Lagrange multiplier α_i^* resulting in

$$fitness_{regression} = \left(R^2 + 1/C\right) \sum_{i=1}^{n} (\alpha_i + \alpha_i^*). \tag{45}$$

3.4 Algorithm implementation

All code written for this thesis was implemented in MATLAB, and any PLS analysis of Raman spectra was performed using built-in functionality. For model generation all datasets were split randomly into a training set (40%) and a validation set (60%). Error bars relating to the output of SVM models are generated using 10 random divisions of training/validation sets in order to account for model variance.

3.4.1 Raman spectral pre-processing

Spectral preprocessing commonly consists of two main steps; normalization and background subtraction. Normalization simply consists of dividing each spectrum by its maximum intensity in order to have data scaled between zero and one. The background subtraction method used in this thesis is the algorithm described by Kandjani *et al* [54].

In short, this algorithm uses a wavelet method to approximate the second derivative of the Raman spectrum. The zero-crossings of the resultant function relate to the edges of a Raman speak. All regions of the spectrum containing peaks are removed, and a polynomial function is fit to the remaining data. This function is then subtracted from the original data to remove the background.

3.4.2 SVM optimization solver

The SVM optimization problems were solved using the sequential minimal optimization (SMO) algorithm proposed by John Platt [72]. The classification problems were solved as per Platt's original work, and the regression problems were solved using the reformulation of SMO by Smola and Scholkopf [70].

3.4.3 Genetic Algorithm

The classical genetic algorithm employed in this thesis proceeded as follows:

1. An initial population of size N_p was generated, wherein each member has a random set of hyperparameter values. Hyperparameters were multiplied by 10^4 before being converted into binary in order to maintain four decimal points of accuracy.
2. Tournament selection was used to find a mating pair. In this process, two subsets of size N_t are selected without replacement, and the fittest member of each subset are chosen as mates.
3. The genome of each of the tournament winners are split at a random crossover point. This is analogous to the formation of haploid reproductive cells. Possible crossover points were restricted in order to keep the original "genes" (hyperparameter values) intact.
4. The genomes are then split at the crossover point and halves are swapped to generate two new complete genomes for offspring.
5. Mutation occurs in these new genomes at rate $r_m \in [0,1]$, which is the probability of a binary value being inverted.
6. Steps 1-5 are repeated until a generation of offspring of size N_p is created.

7. The parent and offspring populations are then combined, and sorted in order of fitness. The half with the best fitness proceed to the next iteration, and the other half is removed.
8. The algorithm halts once the best fitness in the population is four magnitudes lower than the best fitness in the original random population.

There are a number of tuning parameters shown in the algorithm above, namely N_p, N_t, and r_m. Default values for each were used in most experiments. N_p was set to either 50 or 100, N_t was set to 10, and r_m was set to 0.01 or 0.05 depending on the dataset. The choice of tuning parameter resulted from early trial and error without rigorous optimization, the exception being the experiment shown in Section 3.7.

3.5 Early validation of GA-SVM on standard data sets

In order to validate the functionality of the SVM and SVR-GA algorithms, they were first applied to standard data before embarking on studies with Raman spectra. Two sets were applied to this end. The first was a spiral classification problem, wherein the goal is to differentiate between three arms of concentric spirals. The second was the 3D non-linear regression problems developed by Hwang *et al* [73]. Specifically, the "harmonic function" and the "complicated interaction function" from this paper were used. These are given by

$$g_{harmonic}(x_1, x_2) = 42.659\left(\frac{2+x_1}{20} + Re\left(x_1 + jx_2 - 0.5(1-j)\right)\right), \text{ and} \quad (46)$$

$$g_{interaction}(x_1, x_2) = 1.9(1.35 + e^{x_1}\sin(13(x_1 - 0.6)^2)\, e^{x_2}\sin(7x_2)) \quad (47)$$

respectively. The surface plot of each of these functions is also shown in Figure 8C and D. Experiments with these regression problems were performed as per the original work, wherein models were generated using 225 random data points and validated using 10,000 evenly spaced points. The training data was generated with or with noise, yielding an SNR of 4 in the noisy case.

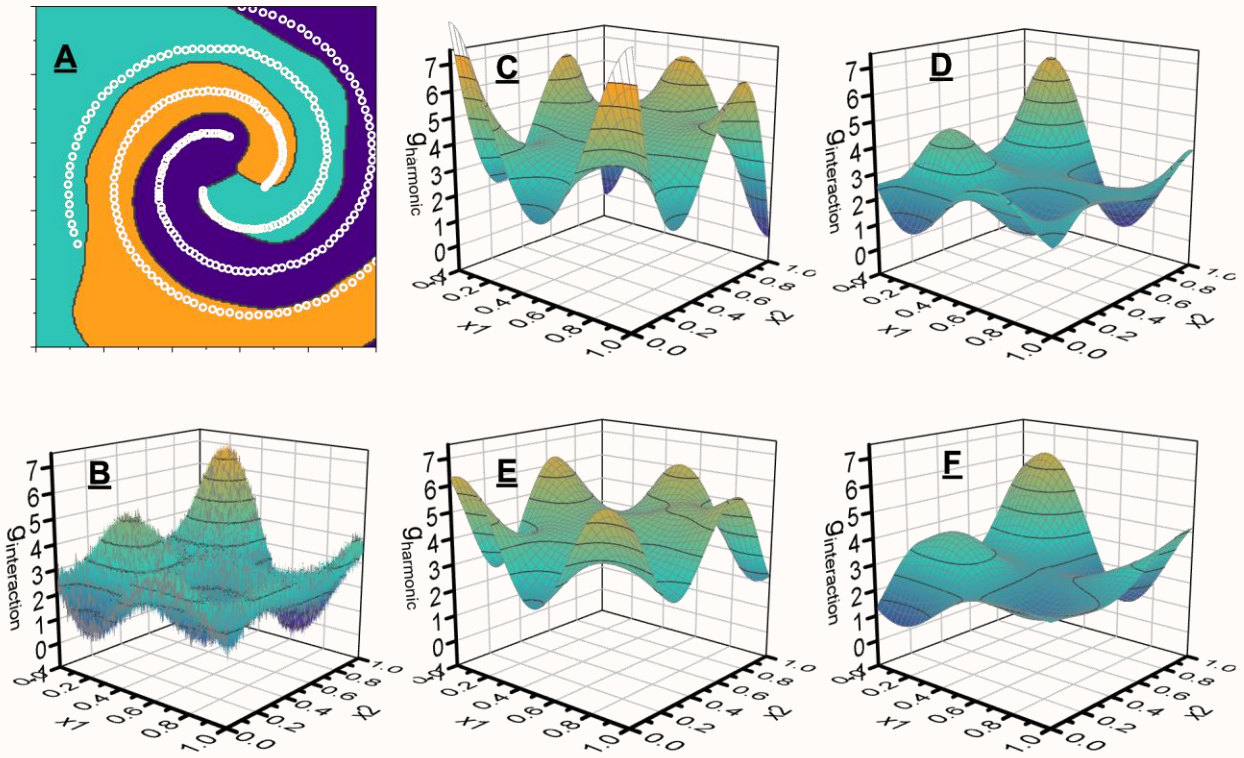

Figure 8: Standard test sets for machine learning. (A) Spiral classification problem, points represent spiral arms to be discriminated, shaded regions are GA-SVM learned spirals. (B) Example of noisy data used to train GA-SVR. (C) $g_{harmonic}$ actual surface and (E) the GA-SVR learned surface. (D) $g_{interaction}$ actual surface and (F) the GA-SVR learned surface.

Table 2: FVU results for different test functions and kernels

	κ_{poly}	κ_{RBF}	κ_{combi}
$g_{harmonic}, SNR = \infty$	1.0752	0.0959	0.0886
$g_{interaction}, SNR = \infty$	0.7914	0.0659	0.0470
$g_{harmonic}, SNR = 4$	0.9932	0.2167	0.1526
$g_{interaction}, SNR = 4$	0.7399	0.2913	0.1524

It is evident that simple linear methods can neither solve the spiral classification problem, nor either of Hwang's functions. Thus SVM-GA was applied using three different kernels; the polynomial, RBF, or combined functions (see equations (37) (38) and (39)). The results of these standard tests are shown in Figure 8A, E, and F. Qualitatively, we can see that the SVM-GA produces expected results for each of these problems. The decisions surfaces for the spiral problem (shaded regions) clearly differentiate the spiral arms, and the surfaces generated for the Hwang regression problems appear similar to the true functions. For the spiral classification GA-SVM achieves an accuracy of 100% for this problem. Quantitative results are shown in Table 2 for the regression problems wherein the kernels are compared on the basis of their fraction of variance unexplained (FVU). This metric is given by

$$FVU = \frac{\sum_{l=1}^{N}\left(\hat{g}(x_l) - g(x_l)\right)^2}{\sum_{l=1}^{N}\left(g(x_l) - \bar{g}(x)\right)^2}. \qquad (48)$$

The first notable result is that in all cases, FVU decreases as the kernel complexity increases, illustrating that the combined kernel seems to perform better than its constituents in all cases. In the case of κ_{RBF} and κ_{combi}, the performance decreases in the presence of noise as one would expect. Unexpectedly however, the performance of κ_{poly} seems to improve slightly upon the addition of noise. This may be a result of the optimization algorithm overfitting the noiseless case, thus reducing the accuracy when new data is introduced. For reference, projection pursuit learning machines tested by Hwang *et al* yielded the lowest FVU values of 0.01045 and 0.04273 for noisy $g_{harmonic}$ and $g_{interaction}$ respectively. These values are certainly lower than those achieved by SVR-GA, but the performance is comparable to these neural networks. It is possible that allowing for additional kernels to be evolved by the genetic algorithm may improve this result. Particularly, using a periodic kernel to learn $g_{harmonic}$ would likely improve the result. However, this would involve prior knowledge about the structure of the data, and it is more common to apply general purpose kernels.

3.6 Validation of GA-SVM on controlled non-linear samples

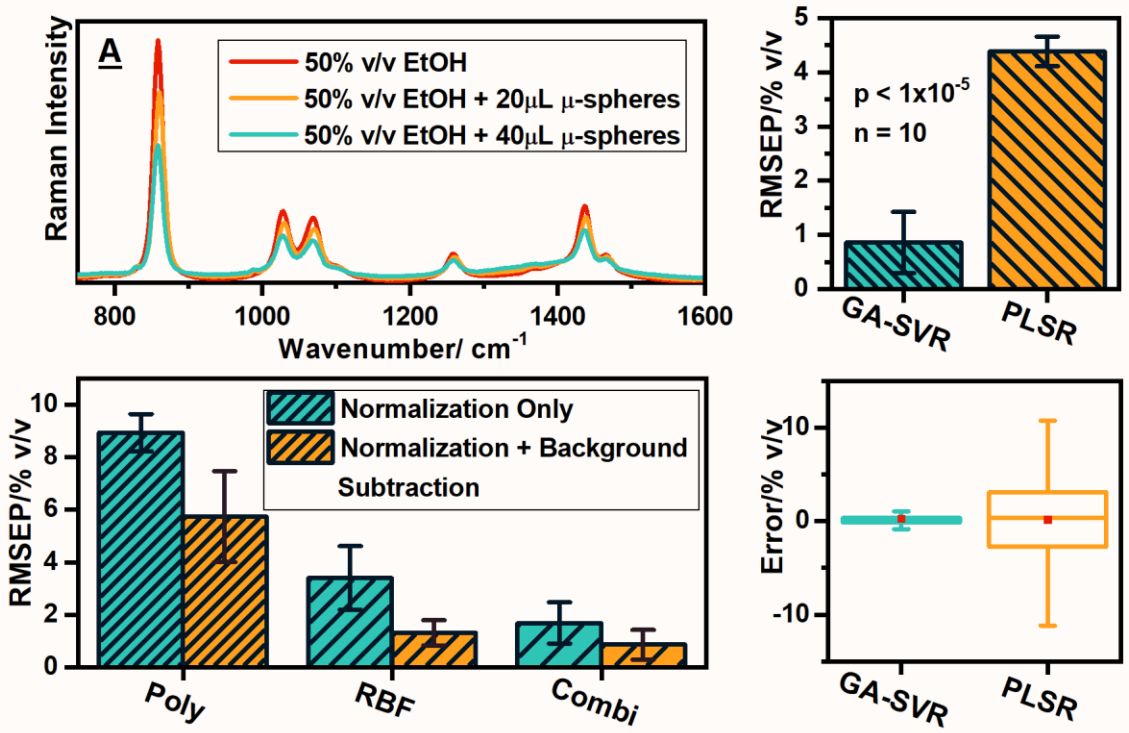

Figure 9: Non-linear regression using Raman spectra of polystyrene doped ethanol. (A) Raman spectra of 50% v/v ethanol with 0, 20, and 40 μL of added polystyrene microspheres. (B) Optimization of kernel type (polynomial, RBF, or combined) and pre-processing methodology (normalization and/or background subtraction) by comparing the RMSEP for different approaches. (C) GA-SVR compared to PLSR on the basis of RMSEP of the concentration of ethanol for 10 random training/validation sets. (D) Box plot comparing GA-SVR and PLSR.

The first real data used to validate the performance of the GA-SVR algorithm was the Raman spectra of different concentration of ethanol (EtOH) to which fluorescent polystyrene microspheres were added. Without microspheres, the Raman scattered intensity is linearly proportional to the concentration of ethanol. When the polystyrene is added, the scattering coefficient is increased thus causing losses of both excitation and Raman intensity. This is illustrated in Figure 9A where the spectrum of 50% EtOH is shown with 0, 20, and 40 μL of polystyrene solution added. The result is that the intensity of the spectrum no longer correlates to the concentration of ethanol in a straight forward linear way. A similar situation arises in biomedical Raman spectroscopy, when differences

in scattering between tissues or cell suspensions causes artifacts in the Raman band intensity that aren't necessarily related to the quantity of certain molecules.

To develop a regression, 0%, 10%, 20%, 35%, and 50% v/v solutions of EtOH in DIW were used. At each concentration, Raman spectra were acquired with 0, 20 or 40 µL of micro-particle suspension, thus yielding polystyrene concentrations of 9.1×10^8 and 1.8×10^9 particles/mL respectively. First, these data were used to optimize the kernel type, and preprocessing algorithm for GA-SVR. It was found that the polynomial kernel yielded an average error of 8.9% v/v which was reduced to 3.4% for the RBF kernel, and further reduced to 1.7% for the combined kernel as shown in Figure 9B. These results, taken with the tests on the Hwang datasets, illustrate how using more complex kernel functions can yield superior results. Furthermore, we have found that removal of the spectral background is also critical in improving performance.

Comparing GA-SVR using the combined kernel and PLSR, is it evident that GA-SVR produces a significantly better regression model. Figure 9C shows that the RMSEP of GA-SVR is about a quarter of that afforded by PLSR ($p < 10^{-5}$). It also appears that GA-SVR is more sensitive to the training data based on the larger error bars in Figure 9C. The random population, tournaments, and mutations involved in the GA algorithm are likely responsible for this. As shown in the box plots in Figure 9D, both regression algorithms are quite accurate with the mean and median error close to zero in both cases, but GA-SVR exhibits significantly better precision. The inner fences of the plot extend about ±1% and ±10% v/v of EtOH from the median for GA-SVR and PLSR respectively. This suggests that GA-SVR develops more robust models for non-linear problems, capable of correctly identifying new data.

3.7 Comparing grid search to GA for hyperparameter optimization

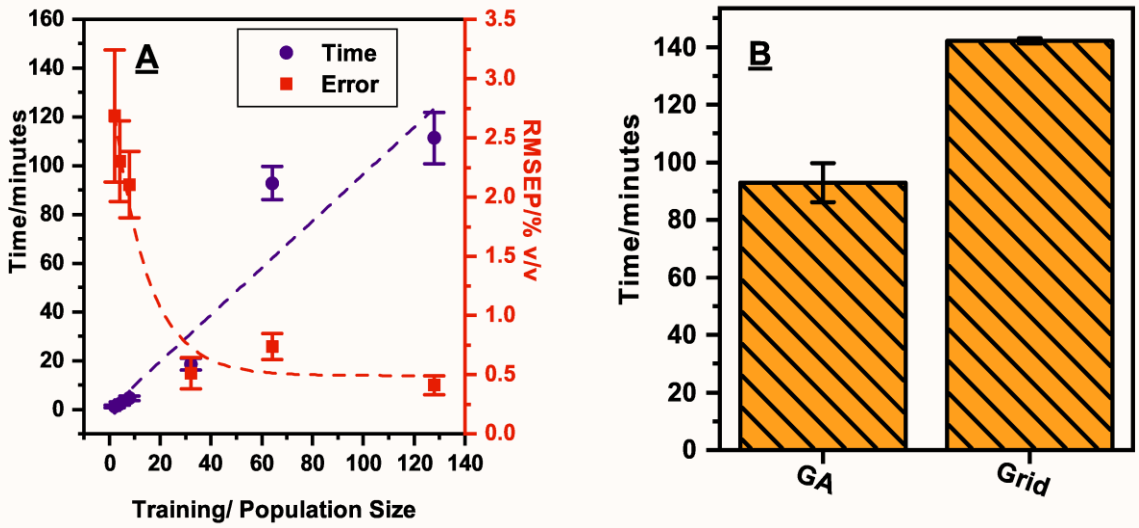

Figure 10: Analysis of the optimization of the SVM hyperparameters with a GA or grid search. (A) Time required and RMSEP as a function of training set and population size for the GA. Lines represent a linear and exponential best fit to the time required and the error respectively. (B) Time required for GA at the optimal population size against the time required for the grid search with two steps per hyperparameter.

It is valuable to understand the time/data requirements for the genetic algorithm to produce a good SVM model. Using the same ethanol/microsphere Raman data and κ_{combi}, the size of the training set was varied from 2 to 200 samples.

Figure 10A shows that the error falls rapidly as the size of the training set is increased from 2 to 50, and saturates to a minimum beyond this point. This suggests that the combined effect of regularization, and the terms in the fitness function that reward robustness, are very effective at avoiding overfitting. Using this minimum training size of 50, the population size for the genetic algorithm was varied from 2 to 140 models.

Figure 10A shows that the optimization time increases roughly linearly ($R^2 \approx 0.9$) as a function of N_p. Larger populations allow for the algorithm to better search the solution space to locate the optimum. This is similar to increasing the grid granularity m as discussed in Section 3.3.5, for which computation time increases exponentially. For comparison, GA optimization with $N_p = 50$ requires approximately 90 minutes to

complete. Using a grid search optimization with $n = 10$ and $m = 2$ requires 140 minutes for these same data. It is therefore evident that grid search optimization requires much more time than the GA approach proposed here, even with an extremely coarse grid size. If one considers the number of times each algorithm must solve the SVM optimization this does not come at much of a surprise. As noted earlier, grid search must run the SMO algorithm nN^m times, or 490 times in the present example. Whereas the genetic algorithm only solves the SMO $N_p N_g$ times, where N_g is the number of generations required to find an optimum. The value of N_g varies run-to-run due to the randomness of the algorithms, but is usually between 5-10. Thus in the present example GA solves the SMO only ~400 times on average, with much superior search capabilities than grid search with $m = 2$. It should be noted that regression model generation with PLS is significantly faster than SVM with either optimization method. This is the cost of the additional precision afforded by GA-SVR over PLS. However, both of these algorithms take very little time to process new samples. So once the models have been generated, computation time begins to become less of a factor.

3.8 Validation of GA-SVM using biomedical Raman data

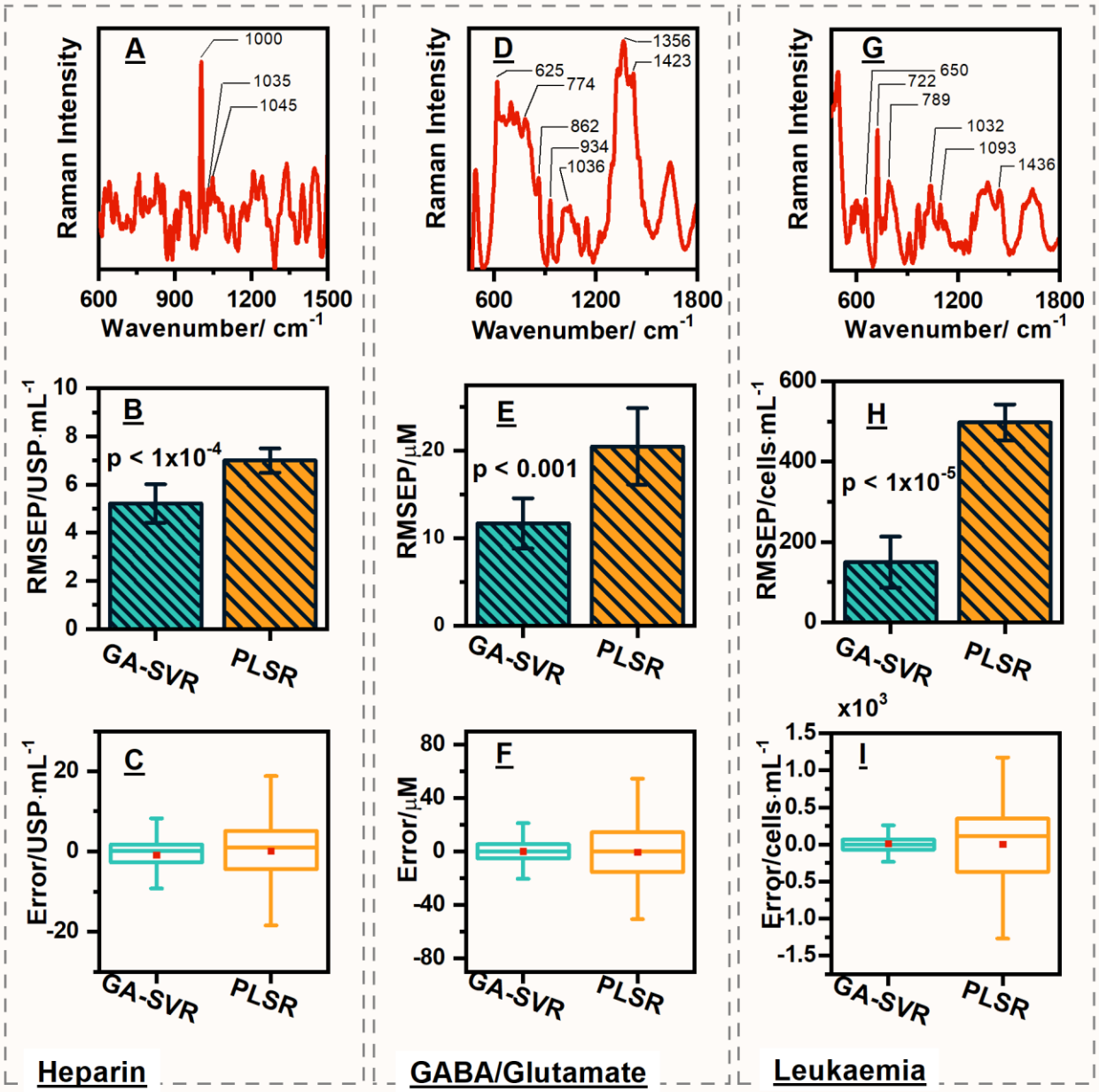

Figure 11: Summary of results comparing GA-SVR to PLSR for biomedical Raman regression problems. (A) Example spectrum of heparin in serum. (B) The RMSEP for predicting the concentration of heparin. (C) Box plot showing the distribution of errors for the predictions of heparin in serum. (D) Example spectrum of GABA and glutamate in serum. (E) The RMSEP for predicting the concentrations of mixtures of GABA and glutamate. (F) Box plot showing the distribution of errors for the GABA and glutamate predictions. (G) Example spectrum for an aqueous suspension of leukaemia cells. (H) The RMSEP for predicting cell density of leukaemia cells. (I) Box plot showing the distribution of errors for predicting the cell density of leukaemia cells.

Next, the GA-SVR algorithm with κ_{combi} was applied to regression problems with biomedical Raman spectra representing several sensing applications. The anti-coagulant heparin is frequently used as a blood thinner after surgery to prevent clots, and careful monitoring is required to avoid side effects[74]. GABA and glutamate are two major neurotransmitters in the brain, and measuring their relative concentrations can yield diagnostic information for a number of neurological disorders including Alzheimer's and Parkinson's diseases[75,76]. Leukaemia is the neoplastic production of pre-leukocytes, resulting in disease due to the reduced population of other blood cells[77]. Each of these datasets have different Raman spectra related to their analyte of interest as shown in Figure 11A, D, G. Please refer to the original publications for the vibrational assignments of the labelled Raman bands.

Figure 11 shows the results comparing regression performance between GA-SVR and PLSR. Similar results to those in Section 3.6 were found; GA-SVR increases precision and reduced error in all test cases. The largest relative improvement was observed in the quantification of leukaemia cells as shown in Figure 11H and I. The error was decreased from 500 cells/mL to 150 cells/mL ($p < 10^{-5}$), and the inner fences of the box plot were reduced from ~500 cells/mL to 100 cells/mL. There is actually some similarity between the quantification of cells and the ethanol/microsphere regression problem. In both situations, the variation in scattering coefficient causes attenuation of the Raman spectrum that may not be correlated to the analyte concentration. Cells exhibit strong Mie scattering of infrared light because cells, cell nuclei, and the organelles within, are similar in size to the incident wavelengths[78]. Additionally, vesicles produced by the cells may also contribute to scattering, and the number of vesicles does not necessarily reflect the cell concentration. Thus, the Raman bands associated with the cells would increase in a non-linear fashion, hence the significant improvement afforded by GA-SVR compared to PLSR.

3.9 Conclusions

In summation, it was observed in all four regression problems that GA-SVMs outperform PLS when using Raman spectra to quantify chemical species. This difference is particularly pronounced for regression tasks in which there is inter-sample variance of

the scattering index. This was shown both in a mock solution containing ethanol and polystyrene microspheres, as well as in real biomedical samples wherein leukaemia cells were quantified using Raman spectroscopy. We have further shown that the use of more complex kernels, which may be constructed from other simpler kernels, can yield better models than the standard approach of only using the RBF kernel. In order to optimize the larger number of hyperparameters necessitated by using more complex kernels, a classical genetic algorithm was applied. It was found that the GA is capable of finding optimal solutions much faster than the grid search algorithm.

The case of discrimination between biochemical distinct samples with Raman spectroscopy and GA-SVM shall be explored more in the next chapter, wherein bacteria are identified using a novel SERS flow-cytometry system.

Chapter 4 SERS Flow Cytometry for Infection Detection

4.1 Introduction

Healthcare associated infections (HAI), in contrast to community acquired infections, are bacterial, viral, or fungal diseases originating from a healthcare setting [79]. The risk of infection is increased in these settings due to the many necessarily invasive procedures that are performed. For example, implanted medical devices such as indwelling catheters, breast implants, or orthopaedic devices have improved patient care, but may introduce pathogens into deep body cavities where they otherwise would struggle to reach. This class of infection is broadly known as implant-associated infection (IAI), and has incident rates up to 40% in immunocompromised patients[80,81]. Different types of IAI primarily differ in their route of entry, and thus the samples used to diagnose them. Of particular interest to this thesis are infections related to orthopaedic joint replacements. These occur at a rate of approximately 2% after surgery, and this rate increases substantially for immunocompromised patients or those with cancer. While relatively rare, this figure represents a large number of people, as over 130,000 joint replacements are performed in Canada every year. Treating a single case of prosthetic joint IAI (PJI) incurs an estimated $10,000 in healthcare costs based on unpublished data from The Ottawa Hospital. Further complicating this issue, the increased use in of antibiotics has led to increased prevalence of antibiotic resistant pathogens [82].

Early diagnosis and screening of infections is important for making informed treatment decisions [83]. The gold standard diagnostic method at this time is to culture bacteria acquired from the infection site. For PJI, samples are usually collected from synovial fluid or as tissue biopsies. The process of culturing bacteria can be quite time consuming, exceed up to five days in some circumstances [84]. This is complicated by the fact that some bacteria may not grow using standard culture protocols, resulting in culture-negative infection [85]. These and other challenges lead the sensitivity of this approach to be as low as 55% [84,86]. After a positive culture, pathogens are identified by

microscopy, antibiotic profiles, antibody/antigen assays, or a variety of related tests. The resultant delay in diagnosis leads to the use of broad-spectrum antibiotics which in turn, may lead to the development of antibiotic resistant infections [87]. If PJI is not promptly and effectively treated, then the implanted device is often lost and the infection may lead to further morbidity or mortality. Therefore, there is a pressing need to improve infection diagnosis by reducing the turnaround time and increasing sensitivity.

Some new technological approaches to bacterial screening, including qPCR and biosensors, are increasingly gaining attention as faster and more precise alternative tools. Rather than detecting the bacteria itself, qPCR based techniques indirectly quantify the microbe by using their nucleic acid, necessitating additional components such as primers, probes, and DNA extraction solutions [88]. In PCR, samples are treated with a series of buffer solutions used to isolate DNA. This DNA is then added to a mixture containing free nucleobases, DNA polymerase, as well as primer sequences. These primers are engineered to be complimentary to genes of interest within a sample. During the PCR process, if the target sequence is present in the sample, it will be amplified by repeated polymerase replication. As a result, only a small amount of DNA is needed in the original sample in order to successfully amplify it by PCR. The DNA can then be quantified by UV absorption. The fact that such a small amount of initial DNA can be amplified by PCR means that these assays can suffer from a high false positive rate in some cases due to contamination. One example of a PCR assay for infection diagnosis is the mecC assay which is used to detect methicillin-resistant *S. aureus* (MRSA) [89]. MRSA is one of the most prevalent and problematic infectious bacteria, as it is resistant to the entire β-lactam family of antibiotics. Differentiating MRSA strains from susceptible forms of *S. aureus* (MSSA) is therefore very important.

Others have reported the use of biosensing techniques such as surface plasmon resonance (SPR) [90–93], and electrical impedance spectroscopy (EIS) [94–96]. These techniques are both affinity assays which measure the subtle changes at the surface of a material when a target binds to it. In SPR, binding of the analyte to the surface modulates the local refractive index, thus changing the momentum matching condition for a surface plasmon mode. The plasmonic substrate could be something as simple as a

metal film, or more advanced structures such as gold nano-hole arrays. In any case the sensing principal is essentially the same, but more complex structures are used to enhance sensitivity. For example, Sarcina *et al* developed a SPR sensor for the agricultural pathogen *X. fastidiosa* by linking antibodies to a gold film using NHS/EDC chemistry on mercapto-R-anoic acid linkers [97]. EIS detects small changes in complex surface impedance when an analyte binds to the surface of an electrode. A frequency swept voltage signal is applied to the electrode, and the Nyquist plot illustrating the relationship between the real and imaginary parts of the electrical impedance is generated. From here, one can correlate aspects of this impedance (magnitude, resistive part, etc.) to the quantity of analyte bound to the surface. As an example, Mondal *et al* designed a disposable impedimetric sensor for bacteria based on glutaraldehyde functionalized filter paper with silver paste electrodes [98]. One notable drawback of both PCR and biosensing methods is the need for specific recognition elements for each target of interest. This makes multiplexed sensors much more complex and expensive.

There have been many SERS based approaches reported for bacterial detection as well. As discussed in Chapter 2, SERS assays can employ intrinsic or extrinsic approaches, which require biorecognition elements in the former. Previously reported extrinsic SERS devices for bacteria include bimetallic nanoparticles with aptamers and Raman reporter molecules for multiplexed bacteria detection [99], and aptamer coated gold and magnetic nanoparticles for *S. aureus* and *S. typhimurium* detection [100]. Examples of the intrinsic approach include a microfluidic system which mixes bacteria and silver nanoparticles for SERS detection [101], and SERS with vancomycin coated silver-gold nanoparticles for detection of bacteria in blood [102]. In many cases it may be necessary to concentrate or isolate bacteria from bulk fluids before SERS measurement. This is because in dilute solutions there is a high probability of no cells being present in the beam waist of the excitation laser. For example, the microfluidic dielectrophoresis enrichment system developed by Wang *et al* was able concentrate bacteria to the point that very dilute concentrations could be measured by SERS [103]. As another case study, Villa *et al* developed a simple method of bacterial isolation by performing SERS on a paper filter [104].

Rather than relying on the magnitude of the Raman spectrum to quantify bacteria, it may be better to count the number of cells that pass through some detection point. This is very similar to flow cytometry, wherein fluorescently labelled cells flow in single file through laser focal points, generating a cell count and fluorescence profile. Raman spectra can be acquired instead of fluorescence to facilitate new applications. For example, dye and antibody conjugated nanoparticles have been used to detect cancer cells in a SERS flow cytometer [105], and a microfluidic trap built over plasmonic nano-dimers could acquire SERS spectra of single cells flowing in blood plasma [106]. Counting of bacteria using antibody and aptamer labelled SERS has also been explored recently with promising results [107].

As discussed in Chapter 2 Section 2.5.2, our group has previously developed a system for acquiring Raman spectra from a hollow-core photonic crystal fiber (HC-PCF) [108], thereby increasing the FOM by having an effective focal length equal to the fiber length. This system is perfectly suited to an optofluidic device such as a flow cytometer, wherein light and fluid co-propagate within the same system. Yang *et al.* have previously utilized a HC-PCF enhanced SERS system to detect bacteria by capillary filling of the fiber and subsequent acquisition of spectra [109]. However, we have previously found that filling by capillary action decreases the robustness of a fiber enhanced Raman system because the filling can be inconsistent, and one then needs to throw away an expensive fiber after the measurement is complete [45]. Our former colleague Altaf Khetani developed a system to reuse these HC-PCFs by repeatedly filling and purging the fiber in an H-shaped flow cell [45]. The resultant design from their work consisted of an HC-PCF bridging two T-junctions into which windows had been cut in order to couple light in and out of the fiber. The flow cell was connected to four reservoirs on each of the other junction nodes. Two were simply waste containers into which spent fluid flowed. The other two reservoirs contained either a sample, or a flushing fluid to clean the fiber. While this system was an integral piece of the development of the optofluidic HC-PCF platform, it was not without issues. Firstly, this system contained a significant amount of dead volume. This is not a significant issue when the sample is plentiful, and there is sufficient fluid to fill both dead and active volume. However, when using samples with less volume, such as those taken from small body cavities like synovial capsules or ovaries,

this may become a problem. Aside from the bulk, there was also issues that arose from the fact that this device was built using components that were not designed for this application. Originally, the system was constructed using high-performance liquid chromatography (HPLC) components made out of polyether ether ketone (PEEK). These components are perfect for the fluidic aspect of the device, as they are biocompatible and can withstand high pressures, they are however less suited to the integration of optical waveguides. The T-junctions were actually made from X-junctions where one of the ports is cut off to make an opening for the optical window. A coverslip is then glued via epoxy to this opening to allow light inside. This process had a lot of room for error, leaks would frequently spring from around the coverslip, and the opacity of PEEK resulted in some practical issues. Since the operator could not see inside the device, they had to couple the fiber without knowing its location, as it was not in the exact same position in the channel every time. Additionally, there was no way to tell if there were bubbles trapped in the system preventing fiber coupling by scattering the laser. All of these issues were addressed by the implementation of a microfluidic flow cell. The micron scale channels in the device reduced the dead volume of the system, while the transparency of the microfluidic material polydimethyl siloxane (PDMS), allowed for easier operation in practice. Furthermore, the flexibility of microfluidic design allows for the implementation of more complex fluidic components that may improve system performance, some of which are discussed in this chapter.

In this chapter, we combine single cell bacterial counting with SERS in an HC-PCF to create a fiber enhanced SERS flow cytometer. Simple silver nanoparticles were used as a SERS substrate in these experiments, and this approach employed intrinsic SERS in order to take advantage of the inherent multiplexing afforded by detecting biochemical differences between cells. This discrimination process was supported by the GA-SVM algorithm developed in Chapter 3. In Section 4.2 and 4.3, the materials and methods are provided so that these experiments may be repeated in the future. These experiments begin with preliminary tests on the silver nanoparticles and microfluidics that supported the subsequent work with bacteria. Next, the performance of HC-PCF enhanced Raman in a microfluidic flow cell, and SERS therein, were explored as discussed in sections 4.4.2 and 4.4.3 respectively. Section 4.5 contains the results and discussion pertaining to

bacterial detection and quantification using fetal bovine serum (FBS) as the biomedical matrix. The bacteria used in this section were *S. aureus* (SA), *P aeruginosa* (PAO1), and *E. coli* (EC). After analysing the system performance, samples containing a single type or mixture bacteria are assayed in sections 4.5.1 and 4.5.2 respectively. The performance of HC-PCF enhanced SERS flow cytometry is then compared to other published approaches to bacterial detection and quantification in section 4.5.3. In Section 4.6, results pertaining to assaying MRSA and MSSA in synovial fluid are discussed. This particular biomedical matrix is of interest for our collaborators in the orthopaedic surgery unit of The Ottawa Hospital. Section 4.7 discusses the preliminary results on functionalizing the HC-PCF with silver nanoparticles to create an all-in-one substrate. Finally, in Section 4.8 several avenues that are currently being explored to improve the HC-PCF flow cell are briefly discussed.

4.2 Materials

Ethanol, silver nitrate salt, sodium citrate tribasic dihydrate, lysogeny broth (LB) medium, agarose, hyaluronic acid, pooled human serum, and rhodamine 6G (R6G) were all purchased from Sigma-Aldrich. Fetal bovine serum (FBS) was acquired from Gibco. Three bacteria species, *Pseudomonas aeruginosa* (PA01), *Staphylococcus aureus* (ATCC 25923*)*, and *Escherichia coli* (CFT073) were acquired from ATCC and cultured following the recommended guidelines. Two other strains of *Staphylococcus aureus* named BP017 and BP038 were methicillin resistant and susceptible respectively. These two strains were clinical isolates obtained from The Ottawa Hospital with approval from the Health Sciences and Sciences Research Ethics Board (file number H-12-18-1994). polydimethyl siloxane with curing agent was Sylgard 184 purchased from Dow Corning. Hollow core photonic crystal fiber used in this work was HC-1550 from Thorlabs.

4.3 Methods

4.3.1 Synthesis and analysis of silver nanoparticles

Silver nanoparticles were prepared using the citrate reduction method. Briefly, 50 mL of a 1.2 mM solution of silver nitrate in deionized water was deoxygenated by bubbling nitrogen gas through the solution for three hours. It was then brought to a boil while stirring

at 700 rpm. Once a boil was obtained, 10 mL of 1% w/v sodium citrate was added. This solution was subsequently boiled for one hr, followed by submersion into ice water in order to stop the reaction. The result was a yellow-grey silver colloid composed of approximately 3.3×10^9 particles/mL with a mean diameter of ~100 nm. This solution was stored away from ambient light at 4°C, and sonicated for one minute prior to each use. The nanoparticles were analysed by means of SERS, transmission electron microscopy (TEM) imaging, and UV-Vis absorption spectroscopy.

Bulk SERS enhancement was calculated as per the conventional formula, which is given by

$$A_{SERS} = \frac{I_{SERS}}{I_{NRS}} \frac{C_{NRS}}{C_{SERS}},$$

where I_{SERS} and I_{NRS} are the measured intensities for SERS and normal Raman scattering respectively. Often measurements for intensity are given in either arbitrary units (AU) or counts/ second (cts/s). AU are used when the spectra have undergone signal processing, and their magnitude is no longer reflective of a physical quantity. C_{SERS} and C_{NRS} are the concentrations of analyte used in the measurement. Often SERS experiments are conducted with a lower concentration so that the amplified spectrum does not saturate the detector. Fiber enhancement is calculated in the same way, wherein fiber measurements replace SERS.

4.3.2 Preparation of synthetic synovial fluid

First a solution of normal saline (0.9%) was prepared by adding NaCl to DIW. Next, this solution along with all necessary tubes and pipette tips were sterilized in an autoclave, and all work was conducted under a flame to avoid contamination of the samples. To 12 mL of saline, which was still 30-40 °C from the autoclave, 45 mg of hyaluronic acid (HA) was added and dissolved under gentle agitation. This HA solution was then divided and diluted in additional saline to produce 6 mL solutions to which 3 mL of pooled human serum was added, resulting in samples that were 1/3 serum by volume. The amount of HA added was chosen to reflect three different joint states; 2.5 mg/mL for healthy samples, 1.5 mg/mL for moderate disease, and 0.75 mg/mL for severe disease. These

solutions were then divided in 0.5 mL aliquots and stored at -20 °C until use. Once thawed, these samples were used immediately and never re-frozen.

4.3.3 Fabrication of optofluidic device

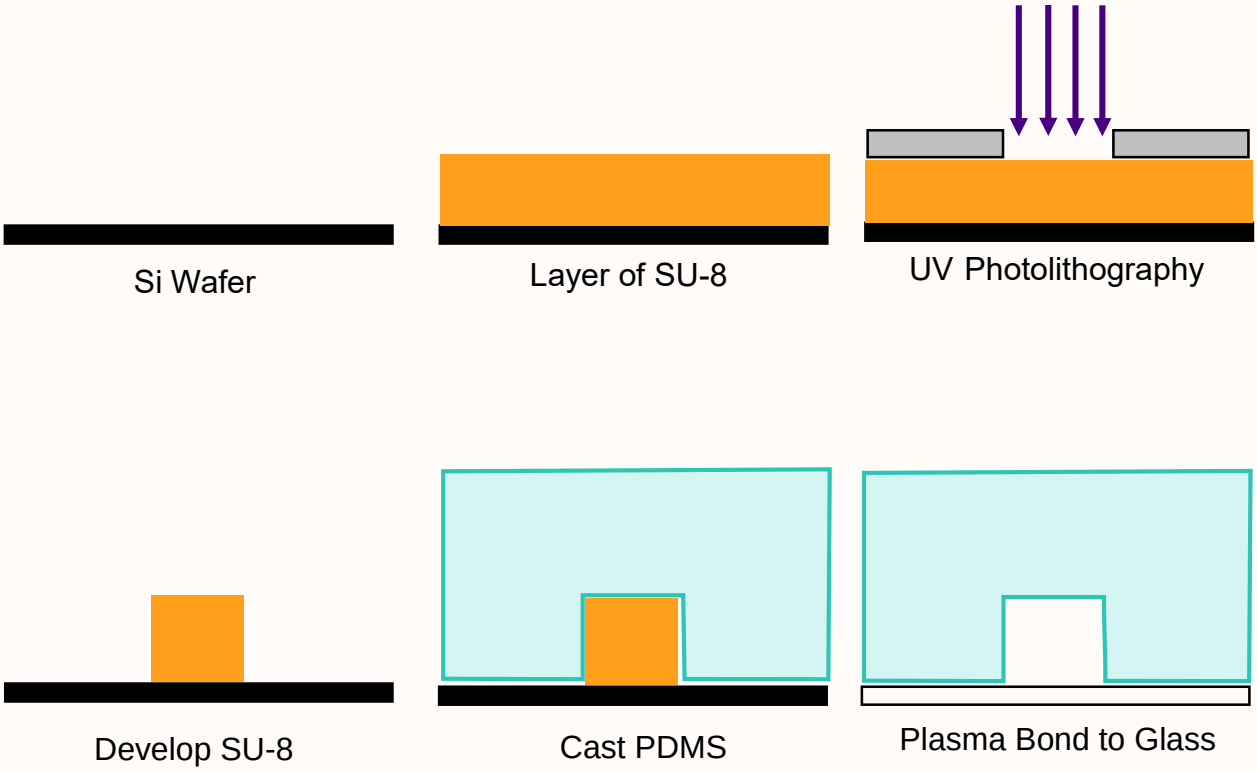

Figure 12: Fabrication process for microfluidic devices base on a simple SU-8 mold

The microfluidic device used in this thesis had two slightly different versions. The first V-shaped design was used in the experiments detailed in Sections 4.4 and 4.5, the later L-shaped design was used for subsequent synovial fluid experiments in Section 4.6. The shift from V to L shape was to better avoid the formation of bubbles in the system which interfered with the HC-PCF. In the V-shaped design, bubbles would get trapped at the top of the fiber interface hole due to buoyancy effects. These same effects allowed any bubbles that formed to float up out of the chip in the L-shaped design, because the outlet channel was positioned parallel to gravity. In either case, the replica molding fabrication approach was the same.

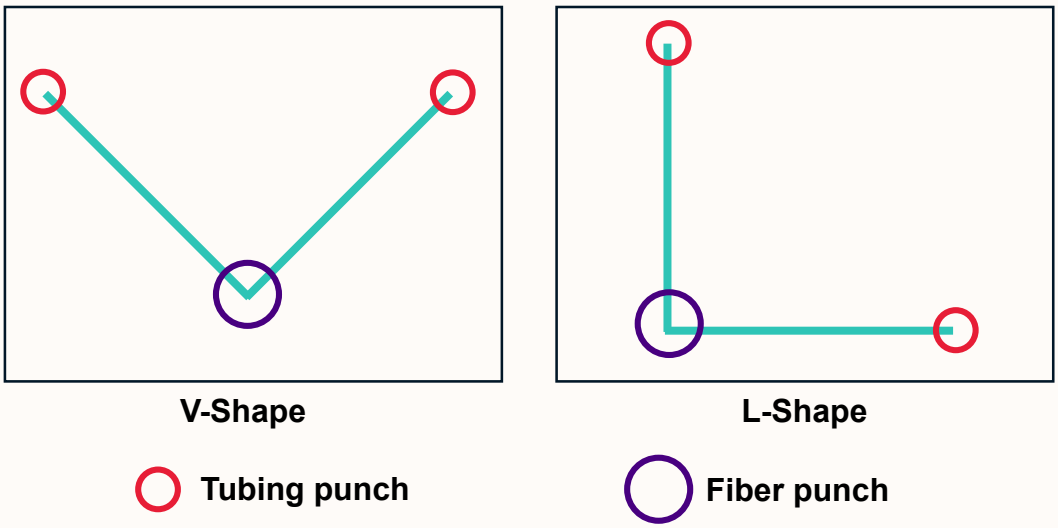

Figure 13: Illustration of V and L shaped microfluidic configurations.

The fabrication initially took place in our collaborator Dr. Michel Godin's lab, and eventually was performed in the Centre for Research in Photonics Nanofabrication Lab. The processes, which is illustrated in Figure 12, started with a silicon wafer which was sequentially cleaned in semiconductor grade acetone, isopropyl alcohol, and DIW before drying with nitrogen. After this, a 60 µm layer of SU-8 100 photoresist was spin-coated onto the wafer using a Laurell WS-650-23 machine. The coating sequence was

1. Ramp to 500 rpm at 100 rpm/s, hold for 5-10 seconds.
2. Ramp to 5300 rpm at 300 rpm/s and hold for 30 seconds.
3. Decelerate to 0 rpm at -500 rpm/s.

The SU-8 layer was then baked according to the manufacturer instructions prior to UV-photolithography. This was performed on a OAI mask aligner using an etched chrome on soda-lime glass mask made by the University of Alberta nanoFAB centre. The SU-8 was exposed to a total UV fluence of 400 $mJ \cdot cm^{-2}$ and developed using standard protocols. The PDMS was then mixed at base/curing agent ratio of 10:1 (w/w) followed by degassing in a vacuum oven for 45 minutes then curing at 70 °C for 2 hrs. The PDMS was demolded and cut to the desired size, then two 0.75 mm holes were punched at the ends of the microchannel for interfacing with polytetrafluoroethylene (PTFE) tubing. Additionally, a 1.25 mm hole was punched in the middle of the microchannel for interfacing with the optical fiber via a sleeve. The PDMS was finally permanently bonded

to a 75×25×1.0 mm microscope glass slide (VWR VistaVision) using air plasma at 50 W for 30 s. For completion of the bond, the device was left at 70°C overnight.

Following the fabrication of the microfluidic devices, the channels were integrated with the HC-PCF to form the final optofluidic device. All components were filled with ethanol during assembly in order to minimize the number of bubbles present in the system. A piece of HC-PCF was used to bridge the two microfluidic channels and affixed with epoxy resin. Finally, the device was flushed with deionized water and was ready to be integrated into the optical system.

4.3.4 Acquisition of fiber enhanced Raman spectra

The system used to acquire Raman spectra was the same as that described in Chapter 2, Section 2.4, except that the sample was an optofluidic flow cell as shown in Figure 15A. The microfluidic device was mounted on a custom 3D printed chassis which was screwed into a Thorlabs 3-axis flexure stage and placed in front of the laser. The flexure stage was used to manually couple the fiber based on optical power meter measurements from the far side of the fiber. The 3D printed chassis restricted the range of possible fiber lengths to 0.5 to 8 cm, which was acceptable for this application. In order to fill the fiber, a 4-reservoir, pneumatically driven flow system was used which is shown in Figure 15. This necessitated each reservoir be connected to an air pressure source (P1-4), which were driven by the building's compressed air system. The air was first passed through an air filter and then controlled by manual regulators. In all experiments P1 was connected to the sample reservoir, P3 was connected to the flushing reservoir, and P2 and P4 were connected to waste reservoirs. The following sequence was used to fill and flush the fiber for measurements.

1. Increase P1 just until the sample begins flowing and allow to fill the microfluidic chip and the tube to P2.
2. Increase P2 until the lateral flow stops.
3. Increase P1 then stop flow with P2 iteratively until both pressures reach 200 kPa and the sample is now flowing through the HC-PCF.
4. Increase P3 to 15 kPa to flow flushing fluid across the tip of the fiber.

5. Once measurements are complete, drop P3 to 0 kPa, and iteratively decrease P1 and P2 to decrease both to 0 kPa while avoiding any back-flow or pressure differences exceeding 50 kPa.
6. Repeat steps 1-5 but reverse the flow so that P3 and P4, both at 200 kPa, are now driving flushing fluid through the fiber to clean it.

The samples containing bacteria were mixed in a 1:1 ratio with the previously synthesized silver colloid. This solution was allowed to incubate while the optofluidic platform was being flushed after the previous measurement, which required 5 minutes. This solution was pumped into the HC-PCF in accordance with the procedure outlined above. The spectrograph was set to acquire one spectrum every second for a period of 10 minutes from the fiber.

4.3.5 Culture and counting of bacteria

Bacteria cultures for quantification were produced by taking 10 µL of bacteria stock suspension and streaking this fluid onto a LB agar plate. The bacteria stocks were solutions of 25% glycerol into which 10^8 – 10^9 CFU/mL of bacteria were suspended before being stored at -80°C. These plates were then incubated upside-down at 37°C for 16 hours. Afterwards, a single colony was taken with an inoculation loop and suspended in 2 mL of LB medium, which was incubated for 16 hours in an orbital shaker at 37°C and 200 rpm. This procedure produced bacteria solutions containing ~10^9 CFU/mL. These sample were then diluted in the bio-fluid of interest (either FBS or synthetic synovial fluid) to create samples with low concentrations of bacteria, which were then stored at 4°C and used within 24 hours for Raman quantification. In order to quantify the samples against the conventional counting method, 5 µL of a 10^5 CFU/mL solution of bacteria was plated onto LB agar plates and incubated at 37°C for 16 hours. The colonies on each plate were then counted in order to get an estimate for the number of CFU in each sample. Mixed samples were prepared in an identical fashion, in that each bacterium was cultured and counted separately before being mixed in solution.

4.3.6 Data analysis

Raman spectra were processed using the methods described in Chapter 3. In this chapter, the rational quadratic kernel (RQK) was explored as the basis for the GA-SVM models which was given in Equation (39). In order to quantify the bacteria, a two-layer classifier was used. The first layer discriminates bacteria spectra from the background serum spectra. The next layer classifies the spectra according to either species or strain depending on the application. This system was used to automatically count the number of bacterial spectral events in a given time frame and then correlating this to the density of cells in the sample. The limit of detection was then calculated based on $LOD = 3\sigma_B$ wherein σ_B is the standard deviation of the blank sample.

4.4 Initial validation of HC-PCF enhanced SERS

4.4.1 Early preliminary experiments

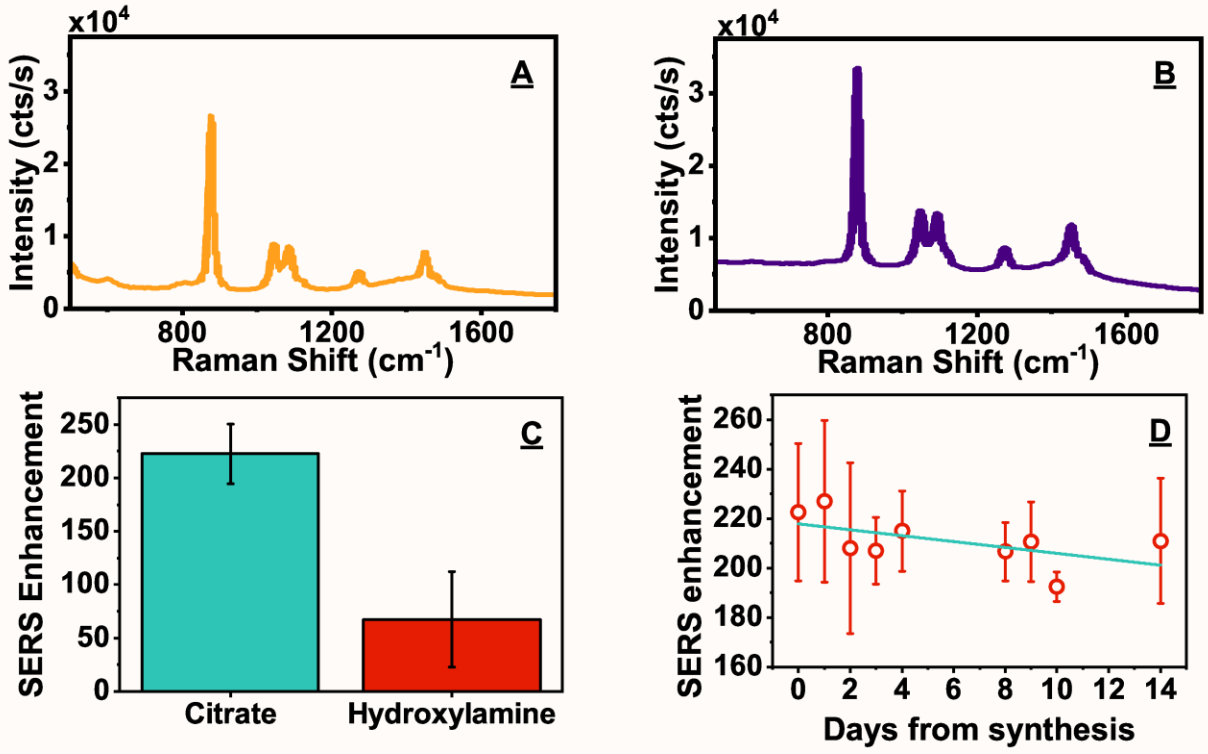

Figure 14: Results of preliminary validation tests. (A) and (B) show the Raman spectrum of ethanol in a 5 cm piece of HC-PCF in an HPLC and microfluidic flow cell respectively. (C) SERS enhancement of citrate synthesized silver nanoparticles versus hydroxylamine reduced particles. (D) Stability of SERS enhancement of citrate-silver nanoparticles over the course of two weeks.

Before proceeding with the development of the optofluidic platform and experimental HC-PCF SERS flow-cytometry, it was first necessary to validate the changes made with respect to previous work. There were two major changes made from previous publications from our group; the switch from HPLC fluidics to custom microfluidics, and a change in nanoparticle synthesis approach. Figure 14A and B show spectra of ethanol in an HC-PCF filled with the HPLC or microfluidic flow cell respectively. The results show that, for the same length of fiber, the two flow cells produce effectively identical Raman enhancements. There is therefore no measurable spectroscopic performance difference when switching to the microfluidic system. Rather, the differences lie in the reduction of dead volume, lower propensity for human error as the need to make an optical window in the HPLC connectors is removed, and anecdotally easier coupling of the HC-PCF. The dead volume in the HPLC system was approximately 100 µL, while for the microfluidic cell this value is halved to 50 µL. This difference seems small, but it is very significant when using low-volume patient samples. It is also substantially less expensive to make PDMS manifolds in-house than purchasing PEEK connectors from suppliers, each of which can cost over $100. Meanwhile, the materials needed to make a PDMS manifold cost only $5, not including capital costs. In terms of nanoparticle synthesis, this work employs a citrate reduction method for the formation of silver nanoparticles as outlined in Section 4.3.1. Previously, our group had used a hydroxylamine hydrochloride based reduction method as developed by Leopold and Lendl [110]. For example, this method was used by our colleague Ali Momenpour for SERS based diagnosis of polycystic ovarian syndrome [111]. Early on in this project we compared this synthesis method to citrate reduction and found that citrate produced a colloid that was both more reproducible and provided a higher SERS enhancement as shown in Figure 14C. We also tested the stability of these citrate silver nanoparticles over a two-week period to observe if there was any significant decrease in SERS enhancement. We found that the particle performance does decay slightly overtime, but the correlation is rather low as shown in Figure 14D. The line showing the linear regression between time and SERS enhancement has an R^2 of approximately 0.2. This indicated that the silver nanoparticles are relatively shelf stable so long as they are stored in the fridge, in an airtight container, and sonicated prior to each use.

4.4.2 Optofluidic Raman platform performance

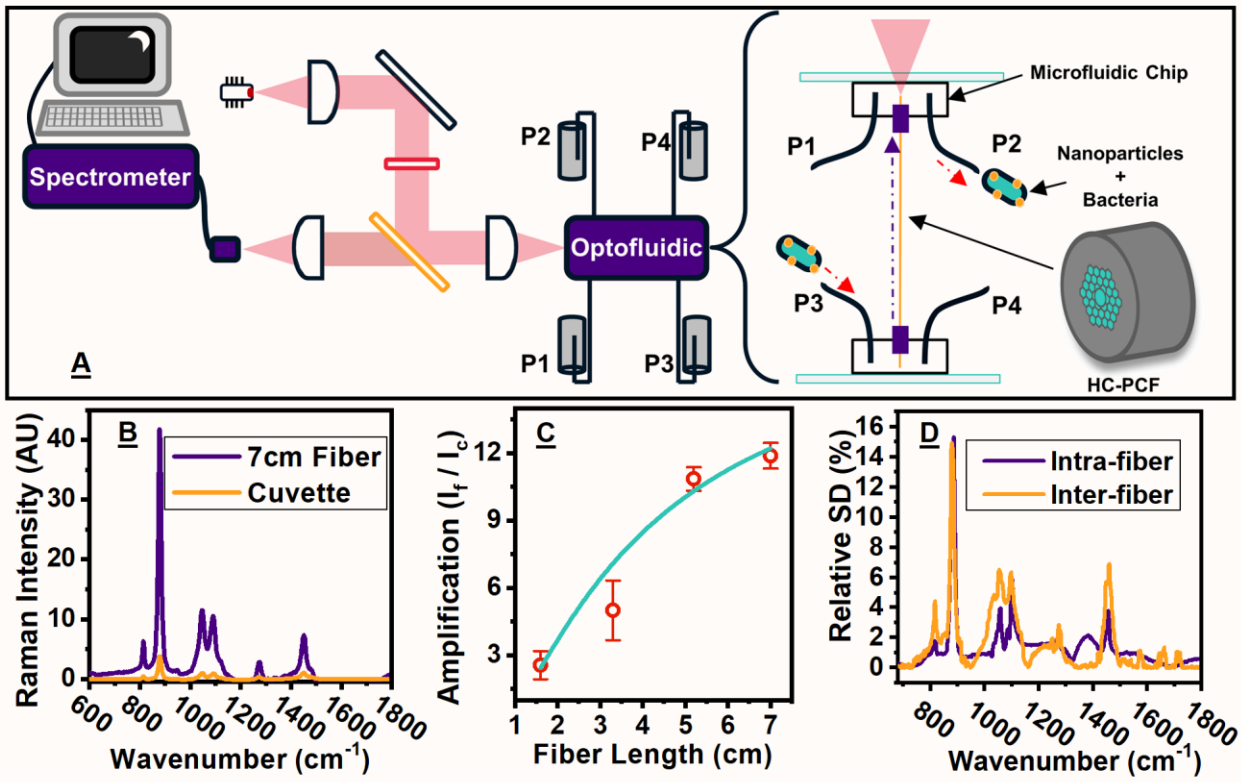

Figure 15: Summary of the general performance results of the Raman optofluidic platform. (A) Drawing of the optofluidic system along with the microfluidic and HC-PCF. (B) Raman spectra of ethanol in cuvette and a 7 cm piece of HC-PCF. (C) Ratio of fiber intensity to cuvette intensity as a function of fiber length, the line is an exponential best fit. (D) Relative standard deviation (SD) of the Raman spectrum measured from 12 fill/flush cycles (intra-fiber deviation), and the deviation between four different fibers (inter-fiber deviation).

The performance of the optofluidic system was evaluated using ethanol to understand fiber length effect on Raman enhancement and the reproducibility of the system. Figure 15B shows the Raman spectrum of pure ethanol as measured from a cuvette compared to that collected from a 7 cm piece of HC-PCF. This fiber yields an approximately twelvefold enhancement based on the peak at 880 cm^{-1}. Figure 15C illustrates the change in this enhancement with respect to fiber length. According to equation (20), there should be a roughly linear relationship between the length of the HC-PCF and the enhancement factor. However, the results suggest that this function is better represented by exponential growth to a steady state. There are two factors that may explain this observation. First, as the length of the fiber increases, the probability of trapping air in the crystal lattice increases. This results in local interruptions in the photonic

bandgap that will decrease coupling efficiency. Secondly, when the fiber is long we must also consider losses of coupled power along its length. Since excitation and collection occurs through the same lens, Raman signal generated at the far end of the fiber will reach the lens less efficiently, and will be excited with slightly less power than signal generated close to the lens.

Figure 15D examines both intra- and inter-fiber reproducibility. These measurements relate to the observed variation in Raman intensity between 12 fill/flush cycles and four different fibers respectively. These metrics range from 5-15% for intra-fiber deviation, and 7-15% for inter-fiber deviation. This result is acceptable for the present application because the spectra are normalized before being processed by the discriminant algorithms. The absolute magnitude of the Raman spectrum isn't strictly relevant because bacteria will be quantified based on the number of spectral events counted as the cells flow through the fiber. If this device were to be used for Raman quantification based on the magnitude of the spectrum, then this reproducibility should be improved. This could be achieved by automating fiber cleaving and coupling to reduce human error.

4.4.3 SERS performance

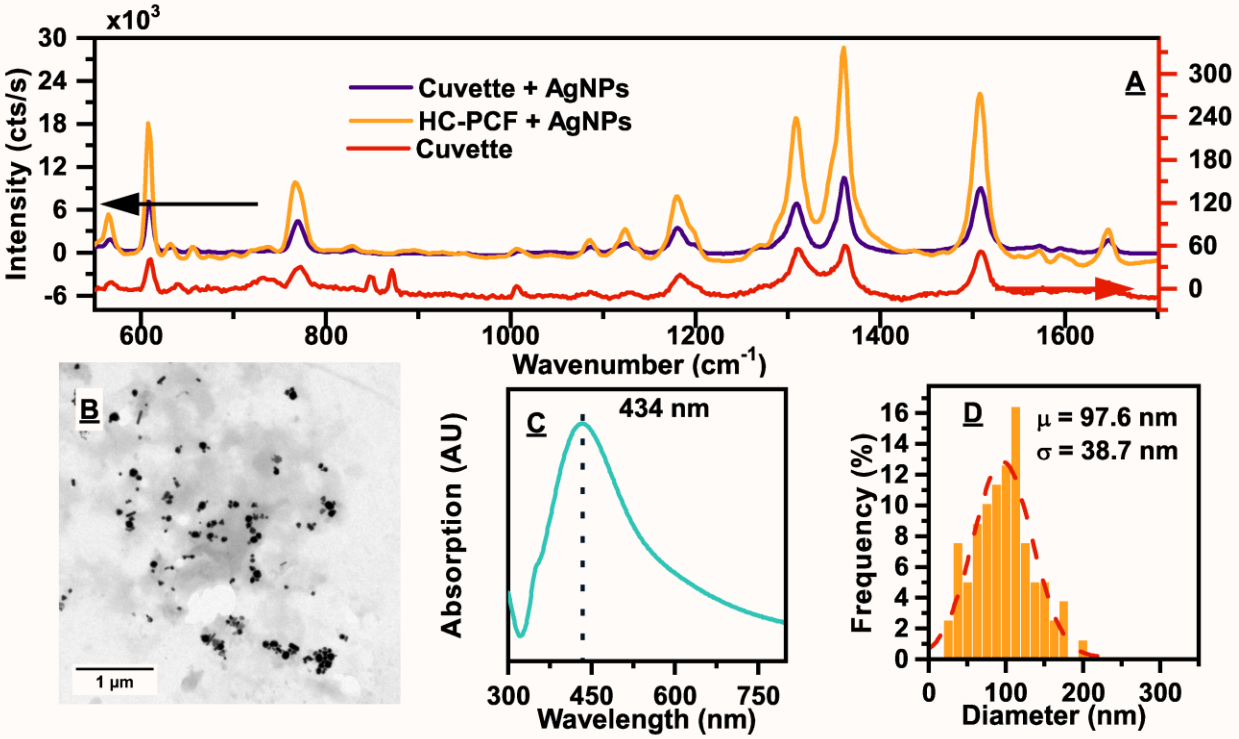

Figure 16: Results of initial SERS validation for the optofluidic Raman platform. (A) Raman spectra of R6G in cuvette (bottom curve and right y-axis), and SERS spectra of R6G in cuvette and HC-PCF (left y-axis). (B) TEM image of the silver nanoparticles on a copper grid. (C) UV-Vis absorption spectrum of the silver nanoparticles. (D) Size distribution of silver nanoparticles.

Next the performance of the synthesized silver colloid was tested, both inside and outside of the HC-PCF. Figure 16A shows the Raman/SERS spectra of 100 μM R6G dye in water in three different conditions. These are normal Raman in a cuvette as a baseline control, cuvette with silver nanoparticles, and finally HC-PCF and silver nanoparticles. The spectrum of the dye in the cuvette is weak and noisy, with a peak-baseline difference of only 60 cts/s for the strongest peak at 1345 cm^{-1}. When the silver nanoparticles are added to the cuvette, the result is a 200-fold enhancement of the Raman spectrum. When this solution is flowed through the HC-PCF via the microfluidic flow cell there is an additional twofold enhancement. The total bulk enhancement of the HC-PCF SERS system is approximately 400 fold over a simple diffraction limited system. The fiber contribution to the total enhancement does not increase if the fiber length is increased beyond 1.5 cm because the coupling of some laser power to the plasmon resonance

modes of the particles attenuates the mode propagating in the HC-PCF. We concluded therefore that 1.5 cm was the maximum effective fiber length, and short pieces no greater in length than 3 cm were used in future experiments. Lengths greater than 1.5 cm were used simply for ease of handling and mounting.

Figure 16B shows a TEM image of the silver nanoparticles. The images show that the colloid has a high polydispersity, and is composed of primarily spherical particles with some nanoplates and rods present. Based on the UV-vis spectra of the particles, as illustrated in Figure 16C, the particles exhibit a broad plasmon resonance band, likely relating to many different coupled modes in the aggregates, with a peak at 434 nm. A 785 nm laser was used to excite Raman scattering due to the high penetration of this wavelength into biological media, and the low associated background fluorescence [112]. However, this wavelength is clearly far from the peak resonant wavelength of the silver nanoparticles. One advantage of the citrate synthesis of silver nanoparticles in this regard is the high resultant polydispersity. This means that there exists several coupled plasmon resonance modes that can be excited by a 785 nm laser. Additionally, we expect that the particles aggregate further and make larger nanostructure upon binding to bacteria cells, although we have not experimentally verified this at this time.

Another important discovery from these experiments was that flushing the system with DIW alone did not regenerate the platform in experiments with R6G. Residual spectrum from the dye could be measured even after flushing. To resolve this issue, DIW was replaced with methanol, into which R6G is significantly more soluble. This allowed the fiber to be fully cleaned between measurements, and eliminated residual signal from the dye. This highlighted the necessity of choosing a flushing liquid that was capable of fully removing any analyte in the HC-PCF to ensure complete system regeneration. When using biological samples, this means choosing a solution that can efficiently remove biofouling from a glass surface.

4.5 Detection and quantification of three different strains of bacteria in fetal bovine serum

4.5.1 Mono-culture bacteria analysis

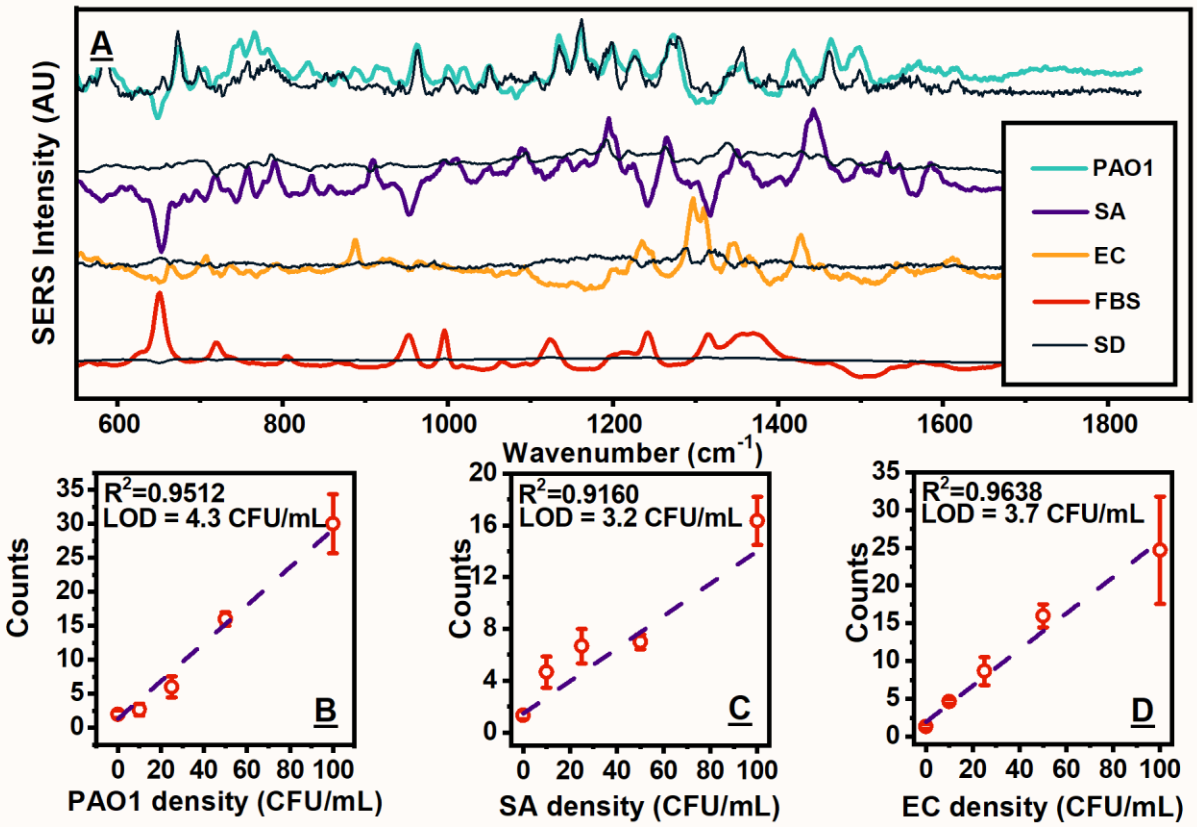

Figure 17: Summary of the detection performance of the device for monocultures of bacteria in fetal bovine serum (FBS). (A) Representative SERS spectra of *P. aeruginosa* (PAO1), *S. aureus* (SA), and *E. coli* (EC), as well as FBS and the relative standard deviation (SD) of each spectrum. (B-D) PAO1, SA, and EC density in FBS as a function of spectral events in a 10 minute window, respectively.

After validating the reproducibility and the amplification of the optofluidic detection platform, the device was used to analyse bacteria present in FBS to simulate clinical samples. As noted earlier, three different bacteria were used for an initial validation of the system, namely PAO1, SA, and EC. Average post-processed SERS spectra for each of these bacteria are shown in Figure 17A, as well as the average spectrum of the background, and the standard deviation of all the spectra. We observe negligible deviation of the FBS spectrum, illustrating that the spectral output from the optofluidic platform is stable between and during measurements. As bacteria cells pass through the

fiber, a very strong SERS signal is generated over a short time-of-flight period as the cell passes through the strong field region in the 1.5 cm closest to the coupling lens. It was hypothesized that the optofluidic platform could be used to count the number of spectral events over a given time frame, and that this count would correlate to the bacterial density in the sample.

The GA-SVM algorithm was used to both automate the counting of spectral events and differentiate between species of bacteria based on their SERS spectra. This first task was performed with a binary model that differentiated the background FBS spectrum from cellular spectra. Using fully processed data with background subtraction made this task relatively easy, because the background spectrum appears as noise after subtraction. Subsequently, the bacterial spectra could be processed with a one versus one multi-discriminant model in order to determine the pathogen identity. GA-SVM was also compared to PLS-DA to determine if SVM yielded similar improvements to classification models as for the regression models in Chapter 3. Spectra were initially acquired from dense monocultures of bacteria to generate a SERS library of ~500 spectra per species. Models were then trained and validated as before, using 60% of the data to train, and the remainder to validate.

The result was that GA-SVM significantly outperforms PLS-DA in differentiating species of bacteria. GA-SVM and PLS-DA yielded accuracies of ~92 and ~56% respectively. The accuracy for simply guessing in a three class problem is 33%, thus PLS-DA does yield some discriminant power but any non-linearities explained by GA-SVM are necessary for a good model. When using simple nanoparticles with minimal surface functionalization as done in these experiments, there is a significant element of randomness in how to particles adhere to the bacteria, as this is a very complex chemical system [113]. Thus there is considerable variation in the SERS spectrum acquired for each species. This is illustrated in the relative standard deviations of the bacterial spectra as shown in Figure 17A. In some cases, the standard deviation is on the same order of magnitude as the average spectral intensity. The result is a substantial degree of within-class variance in the datasets, thus conserved features which differentiate bacteria may be more reliably found in the interaction terms between spectral components. This may

explain why GA-SVM outperformed PLS-DA in this application, because the latter cannot account for these terms in the original feature space. The full confusion matrix showing the GA-SVM performance in building this two-layer discriminant model is given in Table 3.

Samples containing bacterial cell densities between 0 and 100 CFU/mL were then used to validate the quantification performance of SERS flow cytometry with automatic counting by GA-SVM. A spectral event was defined as a change from the background spectrum to a bacterial spectrum and back to the background. Thus consecutive bacterial spectra were considered as one event when calculating the number of cells counted. The reason for this is that the likely time of flight of a cell through the HC-PCF is greater than the spectral integration time in these experiments. If the cells are moving at the maximum fluid velocity, then the Darcy-Weisbach equation estimates that the cells travel at a maximum speed of 1.5 cm/s. This equation is given by

$$Q = \frac{\Delta p D_c^4 \pi}{128 \mu L}, \quad (49)$$

where Q is the volumetric flow rate, D_c is the pipe diameter, p is the differential pressure, μ is the fluid viscosity, and L is the pipe length. As determined by the preliminary HC-PCF SERS experiments, the effective fiber length is limited to 1.5 cm. The integration time of the spectrometer was set to one second for these experiments, thus the fastest a cell could travel through the fiber was in the space of one acquisition. Figure 17B-D show the calibration models obtained for each bacterium using the number of events counted over a period of 10 minutes. Using this system, the limit of detection for bacteria in FBS ranges from 3.2 to 4.3 CFU/mL with good linearity ($R^2 > 0.9$). In order to regenerate the system between measurements, the system was flushed with a 1.0% solution of Triton x100 followed by deionized water. Commercial lysis buffer could also be used as a flushing fluid for this purpose.

Table 3: Confusion matrix showing the results of the dual layer SVM used to count bacterial spectral events and identify pathogens.

		Predicted ID			
		Matrix	**PAO1**	**EC**	**SA**
True ID	**Matrix**	**93.6%**	0.0		
	PAO1		**92.5%**	6.7%	0.8%
	EC	6.4%	4.0%	**90.5%**	5.5%
	SA		1.0%	1.4%	**97.6%**

4.5.2 Mixed culture bacteria analysis

Using the calibration and discriminant models generated in the previous section, inherent multiplexing using SERS flow cytometry was investigated with mixed samples of bacteria. Single blind samples were used in these experiments, which were made by our collaborators in the BeATS lab. The bacterial loads in each sample were not known by the operator until the data had been acquired and fully processed with the GA-SVM. During spectral processing it was assumed that each spectral event represented a single bacterium, and that spectra from multiple species did not temporally overlap. The GA-SVM models were not trained to identify multiple species from a single spectrum, which necessitated this assumption. The results from this experiment are summarized in Table 4. From these results, it is evident that the performance of the device is notably decreased when compared to analysing monocultures of bacteria. The overall root mean squared error is 13.27 CFU/mL, which is most pronounced in blind samples D and E where the load of PAO1 is underestimated, and the load of EC is overestimated respectively. Despite this error, the real and measured quantities of bacteria are highly correlated (R^2 =0.96). This suggests that using spectra from monoculture samples can be applied to the analysis of mixed samples, but some errors are present. One possible contributing factor to this error is the fact that dead cells and cellular fragments are likely registered as

spectral events. The lack of viability metric means that the true CFU of a sample is not measured, because dead cells categorically cannot form colonies. It is possible that cell fragments of different bacteria may have similar SERS spectra, hence the erroneous classification that appears to be observed particularly in samples with higher bacterial loads. It is well known that nano-silver has a bactericidal effect, therefore this may increase the number of dead cells in the sample. It has been shown that the minimum inhibitory concentration of citrate capped silver nanoparticles for EC is approximately 50 µM of total silver. In approximately 1 hour, these particles reduce the number of viable cells by one order of magnitude from a 10^5 CFU/mL starting culture [114]. In this experiment about 60 µM of total silver was used, thus cell death due to the particles was likely. However, the incubation time was very short at 5 minutes, with an additional 10 minutes for the measurement. In previous work by our group, it was found that differences between states of death in leukaemia cells could be distinguished using Raman spectroscopy [66]. Therefore, an additional classifier could potentially be applied using live/dead data in order to get a true CFU measurement, which may improve results.

Table 4: The measured vs. actual bacteria density in five blind samples

Blind Sample	PAO1 (CFU/mL) actual / measured	SA (CFU/mL) actual / measured	EC (CFU/mL) actual / measured
A	10 / 16.64	10 / 7.77	10 / 13.68
B	0 / -2.57	25 / 19.59	50 / 43.23
C	0 / -7.00	0 / -1.09	0 / 4.82
D	25 / 4.82	25 / 29.94	25 / 35.85
E	25 / 15.16	100 / 103.81	50 / 91.99

4.5.3 Comparison to other methods

According to the diagnostic guidelines for PJI, it is recommended that three samples are acquired during surgery for subsequent culturing [115]. Both aerobic and anaerobic cultures are then performed in order to successfully grow pathogens with different metabolic characteristics. These are then incubated for up to 14 days, followed by histopathological analyses. This method achieves sensitivity and specificity on the order of 80% and 90% respectively, and theoretically even a single CFU can be isolated after incubation [115]. Ideally, treatment is delayed until these tests have successfully identified an infecting pathogen so that appropriate antibiotics can be selected. Therefore, it is important to develop new culture-free technologies to reduce or ideally eliminate this two-week lag time.

Table 5 shows the detection limit and time requirements of different techniques reported in the literature compared to the HC-PCF SERS method outlined in this chapter. In all cases, the total assay time is significantly reduced by eliminating the culturing step. However, even in technologies with very low detection limits culturing or sample enrichment may be required in certain applications. This is because isolating even a single CFU from an area of the body that should be sterile can be considered a positive sample. One such example is the case of systemic blood infections, wherein 1 CFU/mL acquired from the blood is indicative of an infection [116]. The benefit of sample enrichment is illustrated in Table 5 by the dielectrophoresis/SERS system developed by Wang *et al* which achieved single CFU/mL detection in 72 hours [103]. The majority of this time was taken up by the dielectrophoresis enrichment process itself, but the result is still faster than culturing some bacteria. Many of the other techniques listed here could be improved by the addition of a cell enrichment method, including our HC-PCF SERS technique. These considerations aside for now, the proposed fiber enhanced SERS flow cytometry approach compares favourably to PCR, SPR, EIS, and other SERS methods. It is worth restating that this is an inherently multiplexed technique by virtue of the GA-SVM discriminants, whereas the PCR and biosensing techniques listed here would require additional primers or recognition elements.

Table 5: Comparison of different novel methods for bacteria sensing including surface plasmon resonance (SPR), electrical impedance spectroscopy (EIS), surface enhanced Raman scattering (SERS), and quantitative polymerase chain reaction (qPCR) assays.

Transducing Element	Limit of Detection (CFU/mL)	Time Required	Ref.
HC-PCF enhanced SERS counting	3.2 - 4.3	15 min.	This work
SPR with gold nanohole array	100	35 min.	[90]
SPR with gold film, anti-fouling layer, and gold nanoparticle sandwich	17 - 7400	80 min.	[91]
SPR with aptamer on gold coated silica nanoparticles	10,000	60 min.	[92]
EIS on gold electrode with antibody	10	Not specified	[94]
EIS on interdigitated gold electrodes with antibody-gold nanoparticles	100	5 min.	[95]
EIS on lectin functionalized gold electrode	75	60 min.	[96]
SERS with dielectrophoresis enrichment and functionalized gold nanorods	1	17 hrs	[103]
SERS with aptamer conjugated silver nanoparticles	15	20 min.	[107]
Duplex droplet digital PCR	4000	70 min.	[117]
qPCR	10	120 min.	[118]

4.6 Assaying MSSA and MRSA in synovial fluid

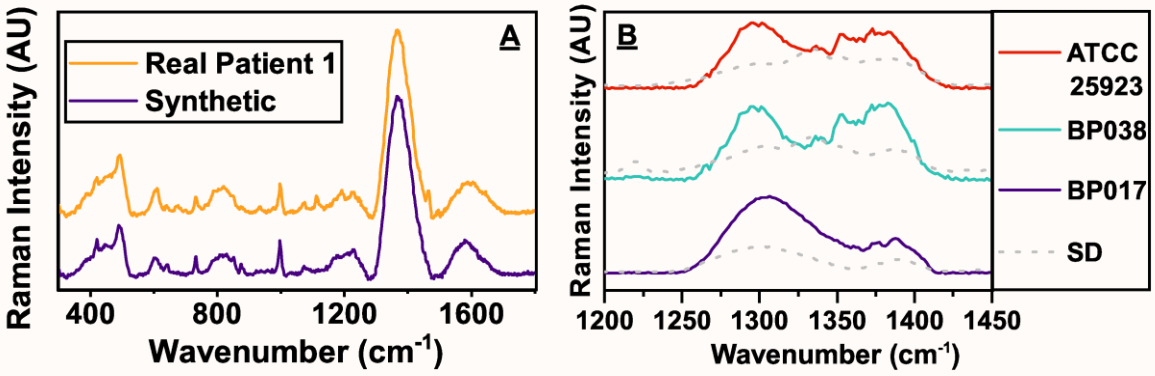

Figure 18: Preliminary results on assaying antibiotic resistant strains of *S. aureus* using the SERS flow cytometry system. (A) Spectra of real and synthetic synovial fluid. (B) Spectra of three strains of *S. aureus* in the amide II region of the SERS spectrum

Following initial validation of the SERS flow cytometry system using FBS as a biomedical matrix, preliminary experiments were conducted in synovial fluid to begin exploring PJI applications. Additionally, as noted previously there is an ever increasing need to rapidly identify antibiotic resistant pathogens in order to appropriately select therapies. The most prevalent of these is MRSA, thus SERS spectra were acquired from three different *S. aureus* strains to determine if they could be spectrally discriminated. These included the ATCC 25923 strain used in the previous sections as well as two clinical isolates from joint infections. These strains, BP017 and BP038, were MRSA and MSSA respectively.

First, SERS spectra in the HC-PCF flow cell were acquired for real patient synovial fluid as well as synthetic fluid made in the lab. The goal was to create a surrogate biomedical matrix in which a spectral library of bacteria could be generated, without needing real patient samples. This necessitates a material with similar chemical composition to real synovial fluid, and therefore an equivalent Raman spectrum. As described in Section 4.3.2, the synthetic synovial fluid is made from a 1/3 dilution of pooled human blood plasma with different amounts of HA to represent different disease states. Synovial fluid is a filtrate of blood plasma containing 20-30 mg/mL of protein, which is roughly 1/3 of the concentration in plasma [119,120]. The three concentrations of HA used are 0.7, 1.5, and 2.5 mg/mL which represent severe disease, moderate pathology,

and normal healthy levels respectively [119,120]. The lower concentrations of HA were not used in the work presented in this thesis but should be used by future researchers continuing this work. Most publications regarding mimetic synovial fluid attempt to replicate its mechanical properties and thus contain other additives. These were not added to the synthetic fluid employed here, because the goal was to mimic the chemical composition rather than mechanical properties. The spectrum of mimetic synovial fluid was compared to the spectrum of fluid acquired from a patient during knee surgery as shown in Figure 18A. The results indicate that the synthetic fluid is a passable chemical replica of the real fluid and can be used to generate a SERS library.

Table 6: Confusion matrix regarding GA-SVM classification of MRSA and MSSA strains.

		Predicted ID	
		MRSA	**MSSA**
True ID	**BP017 (MRSA)**	**93.6%**	6.4%
	BP038 (MSSA)	12.2%	**87.8%**
	ATCC25923 (MSSA)	0.9%	**99.1%**

Next, the spectra of the three *S. aureus* strains were acquired in the synthetic synovial fluid. The average SERS spectra in the amide II band, and associated standard deviations (SD) are shown in Figure 18B. This specific SERS band is shown because this is where the most notable spectral differences between MRSA and MSSA was found. Specifically, we observed an elevated spectral band in the region of 1330 cm^{-1} to 1430 cm^{-1} in the MSSA samples, which is conserved in both BP038 and ATCC25923 relative to BP017. This band is associated with the vibrational modes of CH_3 bonds in lipids, as well as several amino acid modes. One major difference between MRSA and MSSA is in penicillin-binding protein 2A, which is located in the bacterial cell wall [121]. Therefore,

one possible explanation of this observation is that the penicillin binding protein expressed by MSSA shows a stronger binding affinity towards the silver nanoparticles than the proteins expressed by MRSA. This may explain the elevated SERS band that appears in both of the MSSA strains tested.

These spectra were then analysed using the RQK kernel GA-SVM as previously performed, except that the discriminant model was trained using only spectra from BP017 and BP038, and the ATCC strain was left only for the validation set. The reason for this was to determine whether there are indeed conserved spectral differences between MRSA and MSSA. If the ATCC strain is correctly identified as MSSA, despite the model being naïve to its spectra, then this result further implies conserved spectral features. The GA-SVM discriminant results are shown in the confusion matrix in Table 6. We observed an overall prediction accuracy of 93.5%, indicating that the system is able to discriminate between MRSA and MSSA in PJI isolates. Furthermore, the ATCC strain is correctly classified as MSSA, suggesting that there is some preserved structure between susceptible and resistant strains that is apparent in their SERS spectra.

4.7 SERS Functionalized HC-PCF

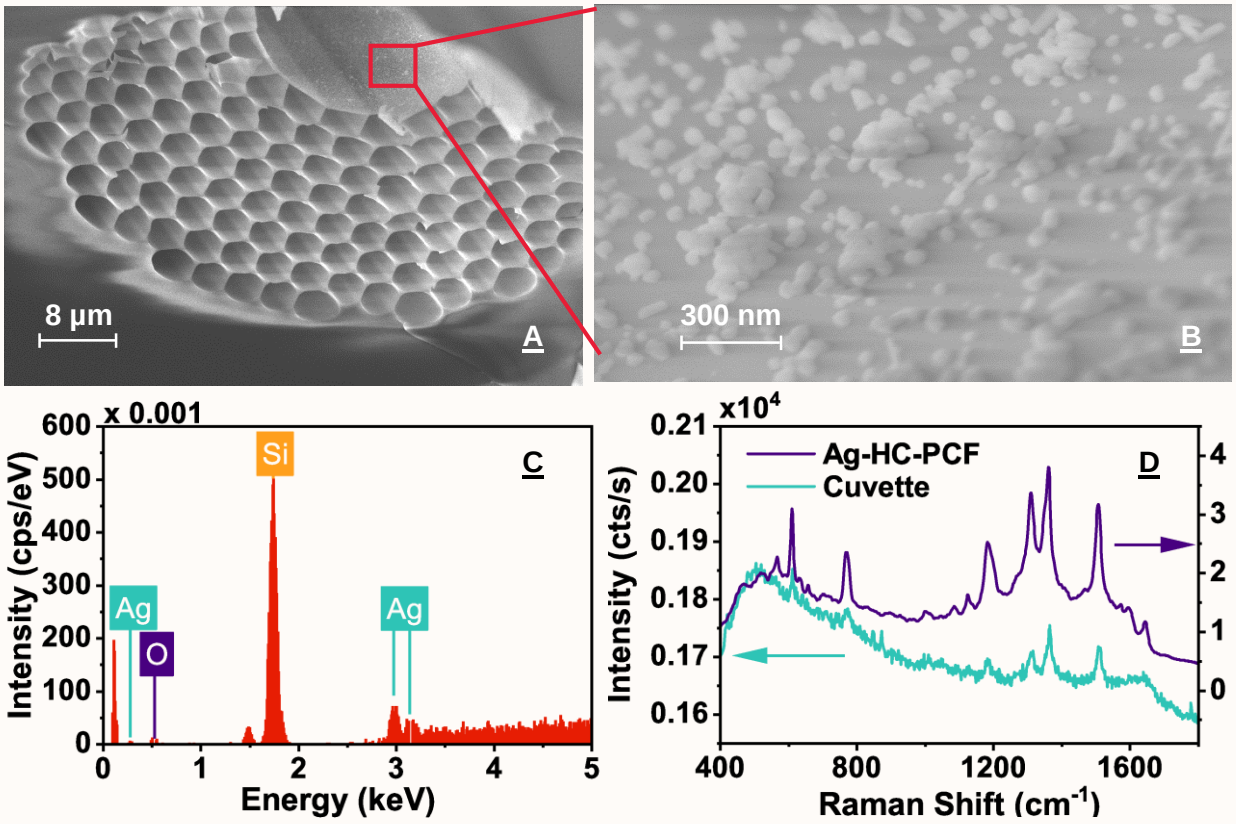

Figure 19: Decoration of the inside of an HC-PCF with silver nanoparticles as a fiber enhanced SERS substrate. (A) SEM image of the cut face of an HC-PCF whose central lattice defect was functionalized with silver nanoparticles. (B) An SEM image showing a magnified section of the inner wall of the HC-PCF. Please note that due to the angle of the fiber relative to the SEM detector, this image does not accurately reflect the particle size and shape. (C) EDS spectrum of the image in (B), with relevant atomic peaks labeled. (D) Normal Raman spectrum of 1mM R6G in a cuvette versus SERS of 0.1 mM R6G in the Ag-HC-PCF

For the sake of posterity, this section will briefly outline a series of experiments attempting to functionalize the inside of an HC-PCF with a SERS substrate in order to build an all-in-one system. While promising results were obtained, the process was difficult and had poor reproducibility, thus more promising research avenues were explored instead of continuing to optimize the fabrication method. However, an interested researcher is certainly encouraged to pick up this work in the future.

Several groups have previously examined the functionalization of micro-structured optical fibers with active plasmonic surfaces. Csaki *et al* examined the deposition of silver using self-assembling monolayer techniques within a suspended core optical fiber [122]. For this application they used the fiber mostly as a sample holder and did not couple light

into it. They showed that the surface layers are stable over several months despite repeated filling of the air-holes in the fiber. Amezcua-Correra *et al* created a microstructured optical fiber with a silver nanoparticle layer generated via chemical vapour deposition [123,124]. In this case, light propagates through the fiber in a solid core and the evanescent field interacts with the plasmonic surfaces created within the air channels. They tested this fiber as a SERS sensor for the detection of ATP and calculated that the over-all enhancement factor of the Raman signal compared to bulk measurement in a cuvette is approximately 10^4. Guo *et al* used a multi-hole capillary modified with gold nanoparticles as a fiber enhanced SERS system [125]. They examined transverse as well as longitudinal measurements of Raman spectra. The study found that longitudinal measurement is superior to transverse measurement which is consistent with our work. Han *et al* compared Raman scattering in solid core PCFs to HC-PCFs with immobilized silver nanoparticles and found that the HC-PCF generated a higher Raman intensity due to the direct overlap between the transmitted field and the sample within [126]. All of these aforementioned studies have found that the transmission efficiency of the fiber is significantly reduced by the introduction of plasmonic surfaces. This leads to a trade-off between increasing the interaction length for Raman scattering and increasing the density of active plasmonic centers to facilitate SERS. The majority of studies focus on forward scattering collection of Raman photons. This means that they have found that it is necessary to use very short pieces of fiber, or a very low silver nanoparticle density. Of the studies mentioned, only one employed backwards collection geometry and found that the trade-off between interaction length and plasmonic surface density actually nullified one another in this set-up [124].

In order to generate a layer of silver nanoparticles on the fiber surface two different routes were explored. The first attempt involved treating the HC-PCF with oxygen plasma to generate reactive Si-O groups on the surface of the glass. Subsequently, a 5% w/v solution of (3-aminopropyl)triethoxysilane (APTES) in ethanol was slowly flowed through the fiber overnight in order to generate amine groups to which the silver nanoparticles would bind. It was discovered that the ATPES had a tendency to gel inside the microstructure of the fiber and prevent flow thereafter, and that the silver nanoparticles also had binding affinity to the plasma treated glass without any need for a binder. SEM

images of the functionalized HC-PCF alongside an EDS spectrum thereof are shown in Figure 19. It should be noted that the odd appearance of the fiber in this image is due to the fiber braking during cleaving. We wished to search for silver nanoparticles in the mid-point of the fiber length, and the plasma treatment process may have caused the fiber to become more brittle. Despite this difficulty, there was evidently silver colloid dispersed over the interior surface of the central defect in the fiber, which could be used to generate SERS spectra. R6G was used as a test molecule to determine the SERS efficiency of this scheme, resulting in the spectra in Figure 19D. The resultant bulk enhancement of the Ag-HC-PCF was found to be 1500 relative to cuvette measurements in these experiments, which is significantly higher than the average enhancement reported in Section 4.4.3 for colloidal silver in the HC-PCF. However, despite these promising initial results, they proved difficult to reproduce.

4.8 Improvements to the HC-PCF flow cell

4.8.1 Microfluidic control system

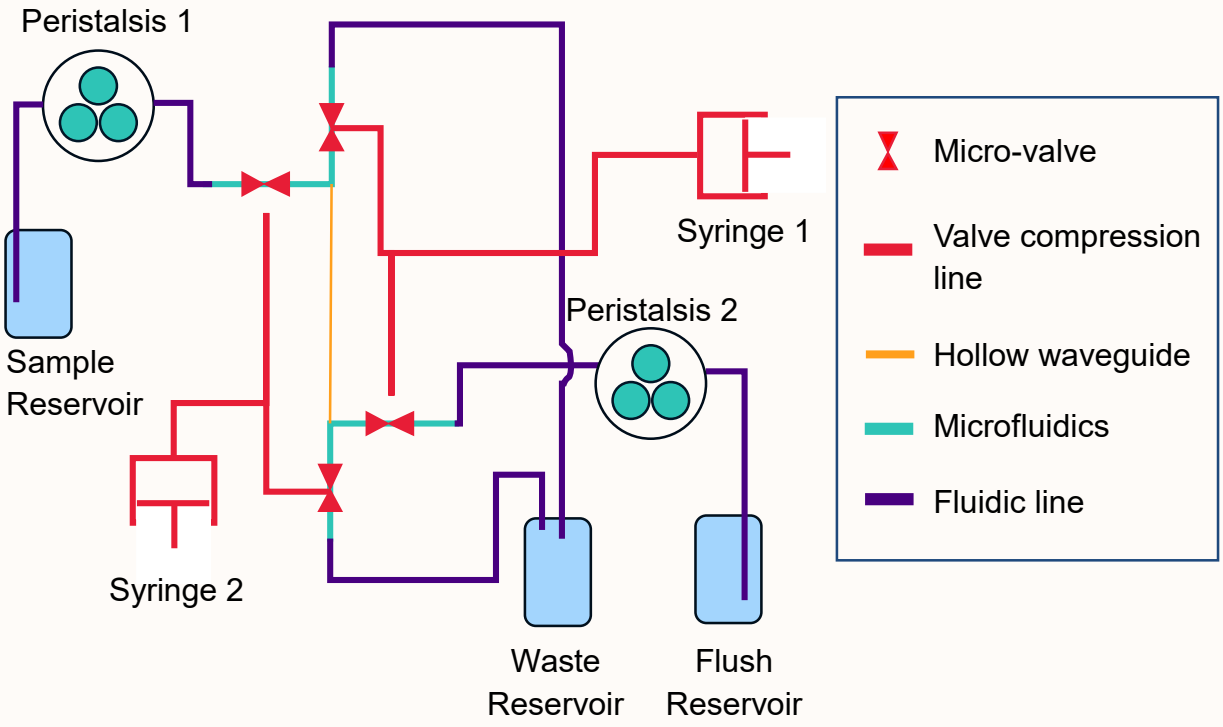

Figure 20: Schematic of the microfluidic control system outlined in this section to potentially replace the device shown in Figure 15. Peristalsis pumps replace the pneumatics, and syringe driven valves are used for precise flow control.

The design for the microfluidic HC-PCF flow cell used thus far is controlled by manual air pressure regulators which drive fluid flow by pressurizing fluid reservoirs. There are several drawbacks to this design which must be addressed to deploy this technology. Firstly, using air pressure to drive flow increases the chance of forming bubbles in the flow cell which could impede optofluidic interfacing with the HC-PCF. This is simply a result of Henry's law, which states that the amount of gas dissolved into a fluid is proportional to the local pressure of said gas. This means that air diffuses into the sample and flushing fluid within the pressurized reservoirs and may precipitate as bubbles in the low pressure region at the far side of the HC-PCF. Therefore, it is important to use another method to drive fluid flow though the system. We are currently in the process of testing peristalsis pumps in conjunction with soft tubing as shown in Figure 20. This class of pumps drives fluid flow by squeezing soft tubing with rollers attached to a rotor. Therefore,

there is now no air pressure which could cause gas to diffuse into the system, and the mechanical components of the pump never come into direct contact with the fluids meaning that biocompatibility is possible with appropriate tubing selection.

Another issue with the previous air pressure control method is that the two sources on either side of a microfluidic chip must be balanced to halt the lateral flow and force fluid through the HC-PCF. This is difficult to do manually and can result in accidental backflow of fluid or air into the channels that may damage the flow cell or decouple the fiber and laser. A possible solution to this problem is the addition of micro-valves onto the fluidic manifolds that will allow for easier control of the flow direction through the HC-PCF. There have been several publications regarding the fabrication of valve systems for microfluidics over recent years. One of the most common approaches used to fabricate micro-valves is the pneumatically driven membrane technique [167]. This involves the deflection of an elastic membrane by air pressure, which either occludes or frees a fluidic layer when moved [168,169]. More novel methods of valve actuation involve thermal control of the size of different materials in order to block fluidic channels. One class of these materials are stimuli-responsive hydrogels which collapse in the presence of high temperatures and swell when the temperature decreases. The size of these gels can be controlled using integrated resistive heaters, optical absorption [171], or magnetic fields [172]. Temperature sensitive fluids can also be applied in a similar fashion [173]. Of these methods, pneumatically driven membrane compression seems to the be simplest method. As such we have designed a microfluidic valve layer that fits over our current manifold design. The manifold is now fabricated using a thin spin-coated layer of PDMS, which can be compressed by the valve layer to halt flow in a certain direction. The valves will be driven by syringe pumps to provide the necessary pressure to compress the flow layer. This whole integrated system is illustrated in Figure 20, wherein both syringe and peristalsis pumps could be controlled by a microcontroller to fully automate the HC-PCF filling/flushing algorithm.

In order to make these valves and thin PDMS layers, standard replica molding is applied as described in Section 4.3.3. However, instead of simply pouring PDMS over the mold, the material is spin-coated for the channel layer. The thickness of PDMS layers for

a given spin speed have been studied previously [174]. In order to demold this thin film, it is necessary to add a layer of sacrificial material to release it. To this end, a thin film of PMMA at most a few hundreds of nanometers thick is first applied and subsequently dissolved in acetone to release the PDMS. The valve layer is made as before and bound to the channel layer by only partially curing both components before aligning and pressing them together.

4.8.2 Integrated microfluidic filtration

As mentioned briefly, implementing some type of cellular enrichment system would be beneficial to this bacteria detection scheme because it would effectively lower the detection limit, and facilitate truly culture free detection. Ideally this method should be fast and capable of separating bacteria cells from any larger eukaryotes that might be present in a clinical sample. This could be performed by centrifugation or by passing the fluid through different sized filter, but it would be valuable to perform this filtration in-line with the HC-PCF SERS system and not require separate processing steps.

One possible method of continuous size exclusion filtration using microfluidic devices are inertial separation approaches, in which the balance of forces on a particle in a flow cause it to migrate to a well defined position within the channel. This is commonly done in curved channels, wherein the centripetal forces on the fluid cause it to form flow patterns known as Dean vortices. These take the form of an even number of counter rotating vortices observed in the plane perpendicular to the major flow direction. The shear forces on a particle from these vortices and the wall lift forces cause particles to migrate to size dependant positions in the lateral cross-section of the flow. This then allows for isolation of certain sized particles from specific regions in the channel. The degree of Dean flow is often described by the unitless Dean number De which is given by

$$De = Re\sqrt{\frac{d_h}{2r_c}}, \qquad (50)$$

where Re is the flow Reynold's number, d_h is the hydraulic diameter of the channel, and r_c is its radius of curvature. One major advantage of this approach is that a large volume

of fluid is processed quickly due to the relatively high flow rates employed. Compared to dielectrophoresis or chromatography, the processing time is reduced, but the cells are also concentrated to a lesser extent. This may mean that multiple inertial separation devices in series may be required to achieve a high enough concentration of detection.

One of the most popular approaches to inertial separation is to design continuously curved channels as spirals, and there have been several reports of using these devices in cell separation. Dr. Jongyoon Han's group has extensively studied spiral microfluidic systems to isolate cells, such as circulating tumors or leukocytes, from whole blood [175–178]. There have been a number of studies investigating the performance of these devices in isolating bacteria sized particles as well. Lee *et al* developed a 3D printed microchannel that successfully separated *E. coli* conjugated to antibody functionalized magnetic nanoparticles from milk [179]. Afterwards the *E. coli* could be detected with absorption spectroscopy. Condina *et al* used a spiral microfluidic system with trapezoidal channels to isolate bacteria from yeast in beer samples [181]. The bacteria could then be successfully identified with MALDI-TOF mass spectrometry. An interesting challenge unique to PJI application is the fact that samples are often in synovial fluid, which is non-Newtonian and exhibits shear thinning properties. The micro-scale fluid dynamics of such biological non-Newtonian fluids has been an area of recent study, including inertial focusing of particles within the flow. It has been shown that even in straight channels, the shear thinning behavior at the channel walls causes lateral migration of particles within the flow [183]. This effect combined with curved channels can result in very efficient size-exclusion filtration and concentration [184].

Flow through a spiral channel inertial separation device is driven by the same peristalsis pumps discussed in the previous section, so long as the flow rate is well defined. The enrichment device can empty cell enriched solution directly into the sample reservoir shown in Figure 20 and be fully integrated into the automated fluid processing. This new iteration of the microfluidic flow cell is currently under development, and constitutes a portion of a colleague's future thesis.

4.8.3 Improving HC-PCF robustness

The HC-PCF employed in this thesis for SERS flow cytometry is a well suited waveguide for this application in several aspects. It guides both the excitation and Raman shifted photons with good efficiency over the fingerprint regime of Raman spectra, and the small mode field diameter in the fiber (10 μm) means that there is a high degree of overlap between the excitation field and the bacteria cells. However, there are a number of challenges encountered when using the HC-PCF for this application. Glass is susceptible to biofouling due in part to its hydrophobicity which encourages the binding of proteins, and bacteria cells can get trapped within the crystal lattice of the fiber which may disrupt the guiding modes and halt flow. It may be possible to employ a different type of waveguide, such as a step index fiber, to address some of these issues. Part of the challenge lies in finding a material which has a lower refractive index than water to facilitate total internal reflection. One such material is Teflon AF polymer, which has a refractive index around 1.29 compared to water's 1.33. Teflon AF capillaries have been tested previously as part of Dr. Altaf Khetani's work, using commercially available material. The problem with these capillaries is that the channel through which fluid and light co-propagate is relatively large at approximately 70 μm. For bacterial detection this is a problem, because now the mode field/cell overlap is significantly decreased compared to HC-PCF. Indeed, Teflon fibers were briefly tried during this thesis' work, and it was found that no spectral events could be detected from bacteria flowing through the fiber. It may be possible to use microfabrication techniques to make a hollow Teflon AF polymer waveguide with a smaller guiding channel to facilitate fiber enhanced Raman. This fabrication process may be somewhat difficult to integrate into a larger system because the chemical inertness of Teflon means it does not bind readily to other materials. Of course, Teflon is still highly hydrophobic, so biofouling may still also be an issue.

An alternative approach may be to modify the HC-PCF in order to take advantage of its existing good properties. One such modification could be to introduce a monolayer of hydrophilic material onto the inner surface of the fiber to prevent adhesion by generating a slip layer of water molecules. One possible route to this could be to add a layer of

polyvinyl alcohol (PVA) using the method described by Trantidou *et al* [185]. In this approach, the surface of glass, or PDMS, is made reactive by oxygen plasma treatment. Then a solution of 1% PVA is flowed through the system to deposit a layer of this hydrophilic polymer on the surface. We have already successfully employed this method to hydrophilize the HC-PCF, with early anecdotal results suggesting the fiber is less prone to plugging as a result of this treatment. In order to prevent bacteria, or other types of cells, from becoming trapped in the fiber crystal lattice, the best approach may be to collapse the cladding holes at the ends of the fiber to prevent any fluid from entering them. The result would be that the fiber becomes a step index guided system because the effective index of the air-filled lattice would be lower than that of water. This extra step should only be necessary in the case of cell-containing samples, and other chemical detection assays could continue to use the fiber as a bandgap guided system.

4.9 Conclusion

In summary, we have designed an optofluidic Raman platform which utilized a hollow-core photonic crystal fiber filled via microfluidics as the main transducing element. This system can be regenerated repeatedly, by flushing the system with a fluid into which the analyte of interest is soluble. Addition of silver nanoparticles into the system provides a large bulk enhancement to the Raman scattered field which facilitates measuring the spectra of biological molecules which tend to have low Raman cross-sections. This system has been applied to the multiplexed detection of bacteria in FBS, wherein a novel method of forcing bacteria to flow through the hollow fiber and counting the number of Raman events was developed. This counting was preformed automatically using a powerful genetic SVM algorithm. This system achieved a detection limit on the order of 4 CFU/mL while requiring only minutes to perform a measurement. Promising results were also obtained for bacterial SERS in synthetic synovial fluid, wherein it was discovered that MRSA and MSSA could be differentiated based on their spectra. Further development of this system has the potential to greatly improve standard of care for patients suffering from infection, as well as many other possible applications in food or environmental monitoring.

Chapter 5 Determination of Chemoresistance in Ovarian Cancers by Kinetic SERS on Cysteine-Capped Gold Nanoparticles

5.1 Introduction

Ovarian cancer (OVCA) is a devastating disease that is often diagnosed at a late stage, resulting in it being the most fatal of all gynecological cancers [127]. After diagnosis, surgical removal of the majority of the tumor mass followed by chemotherapy with a platinum-based drug such as cisplatin (CDDP) is the current standard of care [128]. Unfortunately, relapse is extremely common in OVCA cases and this reoccurrence is frequently accompanied by resistance to the chemotherapeutic agent (chemoresistance) [127]. The biological mechanisms by which this resistance is developed broadly fall into two categories; intrinsic and acquired resistance. Intrinsic chemoresistance arises from innate survival factors of the tumour cell including reductions in drug uptake, increased drug efflux, avoidance of apoptosis, and other mechanisms [129]. Acquired chemoresistance results from genetic pressures originating from the chemotherapy, which rewards the expression of survival factors [130].

Accurate diagnosis of chemoresistance is an important step in delivering the best possible therapy, and there are a number of assays that are currently employed. These include culture-based tests using primary tumour cells which are often acquired during initial surgical debulking. Chemoresistance is determined by exposing cells to the chemotherapeutic drug and then assaying survivability, commonly by measuring thymidine uptake, nuclear morphology, or ATP production [131–133]. These tests can be effective at predicting drug responsiveness but acquired resistance may be missed in the initial tumour cells. Another test involves acquiring positron emission tomography (PET) images using radiopharmaceutical platinum complexes [134–136]. Treatment with these radiolabeled drugs means that at least one round of chemotherapy must be completed, thus diagnosis by PET may occur late into treatment [137]. Other approaches include

measuring the mass spectra of tumour cells and analysing their metabolomic profile with machine learning [138], and Raman micro-spectroscopy to analyse biochemical differences between cancer cells [139–142]. A less invasive approach to diagnose chemoresistance is to quantify different biomarkers in serum. This can be achieved by various methods such as Western blots, ELISA assays, RT-PCR, or biosensors. Specifically, measuring CA-125 and HE4 levels in the blood is used to screen ovarian cancer patients. However, the sensitivity and specificity of these tests can be low, particularly in early stage cancers [143,144].

Tumour derived exosomes represent a new potential biomarker for chemoresistance, which can be measured from the tumors themselves or circulating in the blood. Exosomes are nanoscale vesicles, 30-200 nm in diameter, which are exported from cells and play a role in cell-to-cell communication in healthy physiological processes as well as many pathologies [145,146]. They are formed through the internal budding of endosomes, resulting in a multivesicular body which merges with the plasma membrane to expel the nano-vesicles into the extra-cellular fluid. Tumor-derived exosomes have been implicated in the development of chemoresistance through several cell-to-cell communication mechanisms [147–149]. Additionally, CDDP resistant cancer cells have been shown to exhibit a reduced lysosomal compartment, excrete significantly more exosomal-protein, and export more CDDP per exosome than chemosensitive cells [150,151]. This suggests sequestration and excretion of CDDP through the exosomal pathway is a component of chemoresistance, and quantification of exosome derived CDDP represents a potential diagnostic/prognostic biomarker.

Quantification of CDDP is currently non-trivial and employs complex methods such as high performance liquid chromatography coupled to a fluorometric, or mass spectroscopy quantification methods [152,153]. These techniques require expensive instrumentation and highly trained operators, meaning that their reach may be limited to large research hospitals. There have been limited attempts to circumvent this by creating a bespoke and facile sensor for CDDP, with one example being the g-quadruplex DNA fluorometric sensor developed by Jantarat *et al* [154]. Quantification of exosomes is frequently achieved via nanoparticle tracking, dynamic light scattering, or flow cytometry technology

[155]. In these techniques, elastically scattered light or fluorescent emissions are used to measure exosomes. This is reasonably simple and accurate, but fluorescently labelled CDDP would be needed to also quantify the drug. This would be unrealistic from a clinical perspective as labelling CDDP would likely change its pharmacodynamics and thus its treatment efficacy.

In this chapter, the development and early validation of a SERS chemoresistance assay is presented. The sensor measures three characteristics of OVCA derived exosomes, CDDP content, exosome concentration, and the exosome chemical profile simultaneously to generate a multi-factor diagnostic indicative of chemoresistance. This is accomplished with cysteine capped gold nanoparticles (Au-Cys), which are destabilized upon interaction with CDDP. The rate at which the particles aggregate can be measured by a kinetic series of Raman spectra, and this was found to be proportional to the CDDP concentration. Once the particles have finished aggregating, the steady SERS spectrum contains exosome chemical signatures. These spectra can be used to generate both regression and classification models to simultaneously quantify exosomes and classify them based on OVCA subtype and chemoresistance. This sensor was validated using 6 different cell lines (4 chemoresistant and 2 chemosensitive), from two of the major histological subtypes of OVCA, namely endometrioid (EM) and high grade serous (HGS) carcinomas. Section 5.2 concerns the materials and methods used in this work, including the nanoparticle synthesis method. The results and discussion in Section 5.3 begin with characterization of the cysteine capped gold nanoparticles, and validation of the CDDP sensing mechanism. Next, in Section 5.3.2, exosomes are added to the sample solutions to determine their effect on CDDP quantification and develop a GA-SVR model with their intrinsic SERS spectra to quantify the vesicles. These same spectra were used in Section 5.3.3 to chemically differentiate exosomes between OVCA subtype and chemoresistance profile using GA-SVM. All of these metrics are then combined in Section 5.3.4 to arrive at a chemoresistance diagnostic. Finally, in Section 5.3.5, the sensor is used to study the effect of plasma gelsolin (pGSN) expression on chemoresistance.

5.2 Materials and methods

5.2.1 Nanoparticle synthesis

Au-Cys particles were made via ligand exchange with citrate capped gold nanoparticles (Au-Cit), due the formation of Au-S bonds between the metal and amino acid. To make the Au-Cit particles, 40 mL of 0.5 mM solution of tetrachloroauric(III) acid in DI water was brought to a boil, then 242 μL of 30 mM trisodium citrate was added. This solution was boiled for 30 minutes, during which the solution became a ruby red colloid of gold nanoparticles, which was then cooled to room temperature. The particles were then cleaned by centrifugation at 5000 rpm for 30 minutes, aspirating the supernatant, and replenishing it with clean DI water three times. During each cycle, the particles were re-dispersed via a one minute of bath ultrasonication.

In order to make the Au-Cys nanoparticles, the previously synthesized citrate particles were centrifuged again and suspended in 10 mM NaOH. This solution was then deoxygenated by purging with pure nitrogen for one hour. At neutral pH, cysteine is zwitterionic due to the carboxylic acid and amine moieties possessing negative and positive charges respectively. This causes nanoparticles to rapidly aggregate due to electrostatic attraction when cysteine is added to the solution. To avoid this, ligand exchange could be performed under acidic or alkaline conditions to fully protonate the amine or deprotonate the carboxylic acid respectively. Using a weakly alkaline solution was chosen over the acidic route because the concentration of acid required to impart a primarily positive charge to the cysteine is roughly four orders of magnitude higher than the quantity of base needed to yield a negative charge (carboxylic acid pKa = 1.96, amine pKa = 8.18). Finally, L-cysteine was added to yield a 1 μM final concentration. This concentration of cysteine was chosen based on unpublished inductively coupled plasma measurements of CDDP in OVCA derived exosomes. These showed that the maximum concentration of CDDP was on the order of 1 μM. Thus, this concentration of cysteine was chosen so that interactions between the drug and amino acid would have a significant impact on the capping of the gold nanoparticles. The solution was stirred overnight to facilitate binding of the cysteine to the nanoparticle surfaces. The particles were then

cleaned by centrifugation similarly to above, except 10 mM NaOH is used instead of DI water. Particles were stored at 4°C. All chemicals were purchased from Sigma-Aldrich.

Nanoparticle size and concentration was analyzed by dynamic light scattering (DLS) and ζ-potential measurements performed using a Malvern ZetaSizer instrument. For gold nanoparticles, solutions were diluted 10^4-fold in degassed and centrifuged (5000 rpm, 30 minutes) 10 mM NaOH. Exosomes were diluted 10^3-fold in normal saline to avoid multiple scattering events in the solution. Plasmonic properties of the nanoparticles were assayed using UV-Vis spectra acquired with a Bio-Tek Synergy HT 96 well plate reader. Finally, morphology and size of the nanoparticles was determined with TEM images from a FEI Tecnai G2 Spirit Twin TEM analyzed with ImageJ.

5.2.2 Raman measurement

Using the same basic optical setup as previous studies, Raman spectra were acquired from a capillary based sample cell. The glass capillary was held by a 3D printed mount which positioned the sample such that the propagation axis of the laser aligned with the long axis of the capillary. The mount was attached to a 3-axis flexure stage for alignment.

5.2.3 OVCA cell culture and exosome isolation

Chemosensitive (HGS: TOV3041G; EM: A2780s) and resistant (HGS: OV90, TOV3291G; EM: A2780cp; TOV112D) OVCA cells were used in our study. The HGS and EM cell lines were cultured and maintained in OSE (Wisent Inc, St-Bruno, Canada) and RPMI 1640 (Life Technologies, NY, USA) media, respectively[156]. Media were supplemented with 10% FBS (Millipore Sigma; St. Louis, MO), 50 U/mL penicillin, 50 U/mL streptomycin, and 2 mmol/L l-glutamine (Gibco Life Technologies, NY, USA). All experiments were carried out in serum-free media using 1.6 x 10^6 cells. Conditioned media from OVCA cultures was centrifuged (2500 g, 10 minutes) to remove dead cells and debris. The supernatant was then centrifuged (20,000 g, 20 minutes, at 4°C) to remove microparticles. Finally, the supernatant was centrifuged (100,000 g, 90 minutes, 4°C) resulting in a pellet of exosomes. Exosomes were characterized as previously demonstrated in [147]. These are then suspended in isotonic saline for Raman analysis.

All of these steps were performed by our collaborators in Dr. Tsang's group at The Ottawa Hospital Research Institute.

5.2.4 CDDP and exosome SERS assay

Samples containing exosomes were first sonicated for 15 minutes to lyse the vesicles and release any CDDP and exosomal protein. Fifty μL of sample solution was then mixed with 1 mL of nanoparticle colloid and left overnight. Immediately prior to a measurement, the colloid was destabilized by rapidly adding NaCl and then mixing vigorously for 10 seconds. A 3 cm piece of glass capillary (ID = 1mm) was then dipped into the solution and subsequently placed at the focal point of the Raman excitation laser. Raman spectra were then acquired with an integration time of 10 seconds over a period of 15 minutes, resulting in a kinetic series of 90 spectra per sample.

5.2.5 Spectral data processing

As before, spectra were processed by normalization and background subtraction. However, in this work the airPLS algorithm was used to fit the background function to the spectra [157]. For discriminant and regression models, the GA-SVM algorithm was used with the RQK kernel function. For Raman kinetic measurements, the band at 320 cm^{-1} which is related to a Pt-Cl vibration in CDDP was selected to create a time series. This series was then windowed in the region between 1.5 and 15 minutes, and the slope of a linear regression was used to estimate the aggregation rate in cts/minute.

Multifactor diagnostics scores (S) were constructed using a linear combination of the CDDP concentration, exosome concentration, and SVM score given by $S = wf$ where the score f is a vector containing the individual factors and w is a vector of weights. An optimal w is found using a linear SVM which results in $w = \sum_{i=1}^{n} \alpha_i y_i f_i$, wherein $y_i = \{-1,1\}$ and represents the class (chemosensitive or resistant) of a set of measurements for a single sample f_i.

5.3 Results and discussion

5.3.1 Exploratory study and initial validation

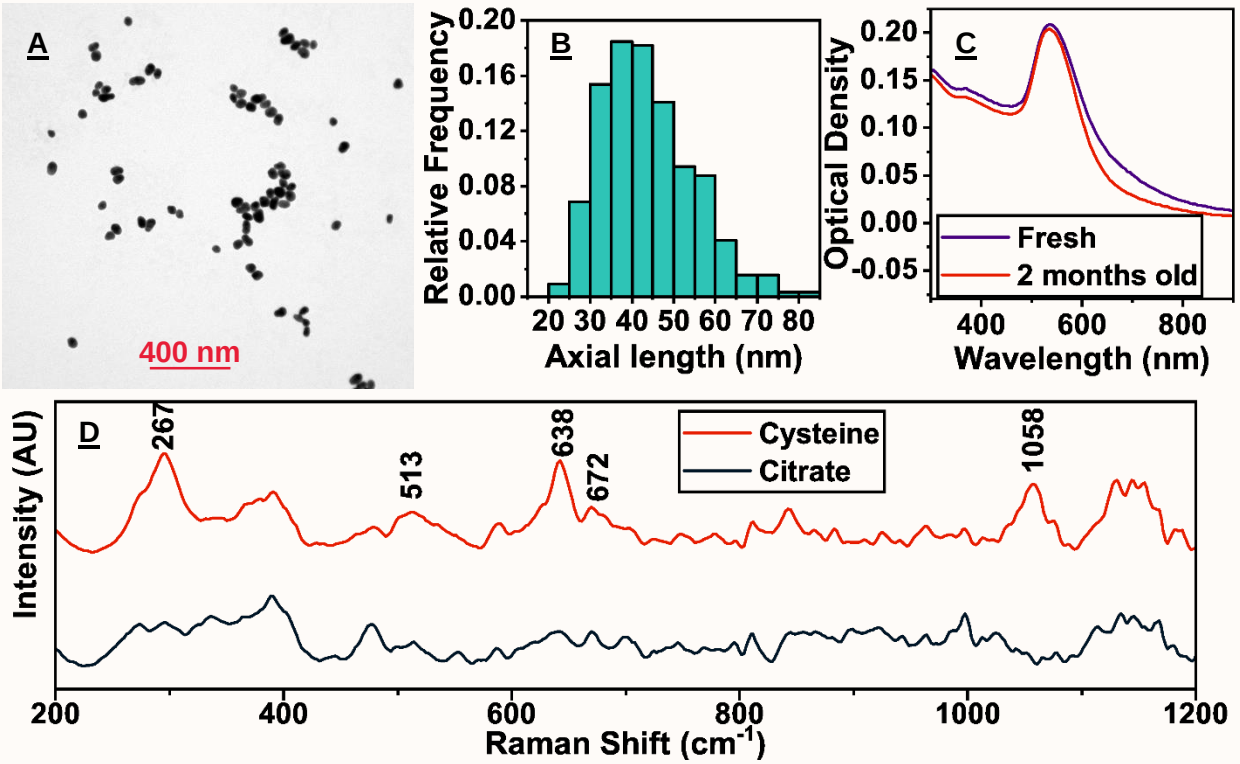

Figure 21: Characterization of the synthesized Au-Cys nanoparticles. (A) TEM image of fresh particles. (B) Size distribution taken from ~300 nanoparticles in different TEM images. (C) Absorption spectra of fresh and two-month old Au-Cys particles. (D) SERS spectra of Au-Cys and Au-Cit particles, with the relevant cysteine/cystine peaks labeled.

Early experiments attempting to use SERS to quantify CDDP relied on directly measuring the bond resonances of the drug. However, it was quickly found that the spectral fingerprints of CDDP were too weak due to the dilute solutions of exosome-derived CDDP. Thus we sought an alternative means of quantifying the drug. One hypothesis was that by capping the nanoparticles with a molecule that would interact with CDDP, the drug would change the stability of the particles. As noted in Chapter 2, when monodisperse particles aggregate their LSPR modes couple together creating plasmonic hotspots between particles. These coupled modes are resonant at longer wavelengths, which is ideal given that Raman scattering was excited at 785 nm in this work. Therefore, one would expect to see the SERS intensity initially rise due to the formation of these nano-scale aggregates, and eventually fall due to the formation of overly large structures. The rate at which this occurs would be a metric for the original stability of the colloid, with

a less stable suspension aggregating more quickly. The SERS bands of the capping molecules or CDDP trapped between particles could be monitored as a function of time, and the change in intensity with respect to time could be used to measure this rate.

For this scheme to function, the nanoparticles needed to be capped with some molecule that would interact or react with CDDP to some degree. The drug normally functions by binding the N7 atom of two adjacent guanine residues to prevent DNA replication [158]. However, CDDP actually binds non-specifically to many nucleophiles such as thiols and amines in proteins. As noted in Section 5.2.1, particles were initially formed by citrate reduction. When ionic gold is reduced to its atomic form by citrate it is oxidized to dicarboxyacetone and acetone. Excess citrate in the reaction then caps the nanoparticles, yielding a negative surface charge on every particle which keeps them suspended by electrostatic repulsion [32]. Citrate is a potential candidate for a capping molecule to interact with CDDP, as its carboxylic acid groups are weakly nucleophilic. However, we wished to introduce stronger nucleophiles to the nanoparticle surface, thus citrate was displaced with cysteine which has both an amine and thiol group.

After synthesis of Au-Cys particles they were characterized by means of TEM, ζ-potential, absorption spectroscopy, and SERS. Based on the TEM images, the colloid contained heterogeneous elliptical particles with a mean axial length of 43.7 nm as shown in Figure 21A and B. To assess the stability and plasmonic character of the particles, absorption spectra were acquired from fresh and two-month old preparations. Figure 21C illustrates these results and shows the expected dipole localized surface plasmon resonance for gold nanoparticles at 533 nm. Furthermore, the small differences between the differently aged particles shows that they are stable for long periods of time. SERS spectra of nanoparticles dried onto a glass slide were used to confirm successful binding of cysteine to the gold surfaces as shown in Figure 21D. The presence of several Raman bands in the spectrum of the Au-Cys nanoparticles suggest that the ligand exchange reaction was successful. Primarily, the Au-S stretch band at 267 cm^{-1} implies that cysteine is bound via the thiol group as predicted [159]. Additionally, strong bands at 638 cm^{-1} (C-N stretch), 672 cm^{-1} (C-S stretch), and 1058 cm^{-1} (H-N-C bend) all imply the presence of cysteine on the gold surface [159]. The band appearing around 513 cm^{-1} is associated

with the S-S bond, suggesting that some cysteine is oxidized during the reaction [159]. The presence of several overlapping bands between the Au-Cys and Au-Cit particles suggests that displacement is not complete during this reaction, and some citrate remains on the particle surface.

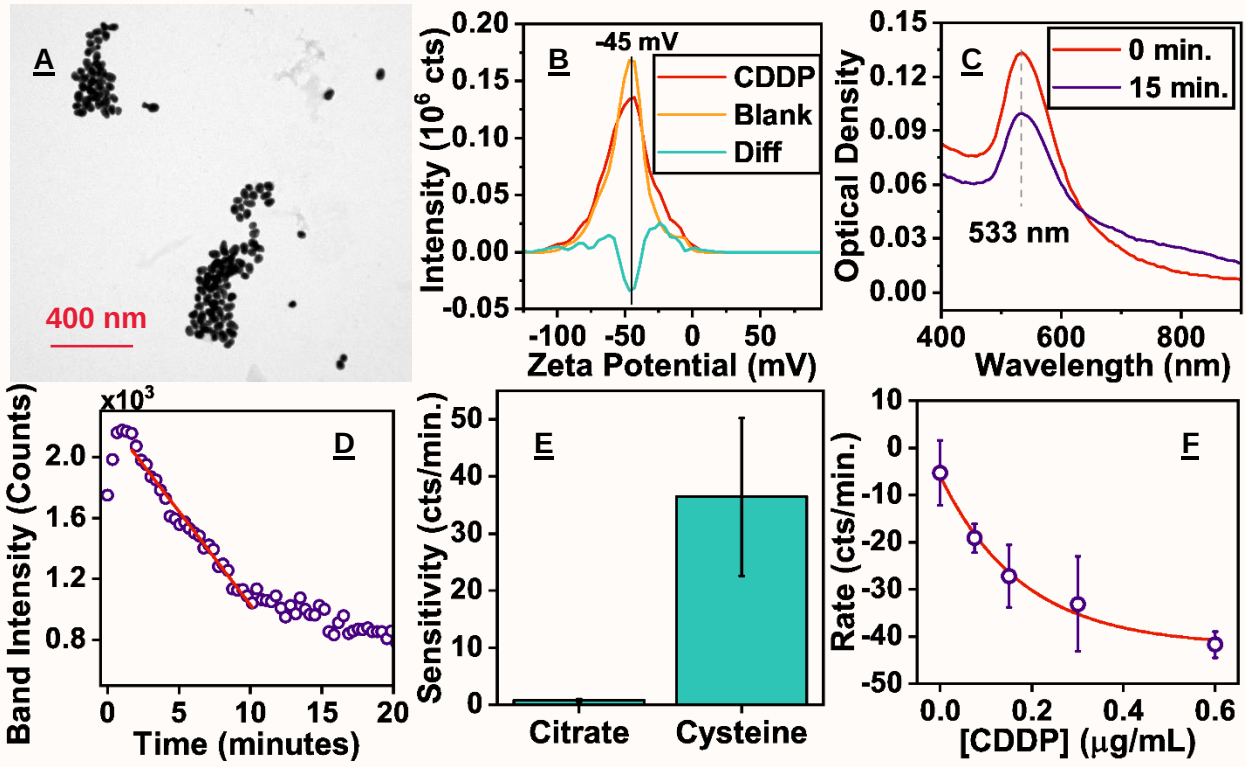

Figure 22: Early validation of the CDDP sensing mechanism using samples of pure drug in isotonic saline. (A) TEM image of the aggregated particles. (B) ζ-potential of the Au-Cys particles with and without interactions with CDDP as well as the difference between these measurements. (C) Absorption spectra of the nanoparticles before and after aggregation. (D) Representative Raman kinetic series, illustrating the progression of spectral intensity over time. (E) Sensitivity of Au-Cit versus Au-Cys particles to CDDP. (F) Relationship between the aggregation rate as measured by kinetic SERS and the CDDP concentration.

To find if CDDP had any effect on the stability of the particles, Au-Cit and Au-Cys particles were incubated with a high concentration (0.6 µg/mL) of drug, with blank isotonic saline and nanoparticles used as a control. To induce aggregation 10 µL of 1 M NaCl was added per 1 mL of nanoparticles to disturb the electrical double layer that normally provides stability. Aggregation was confirmed by the TEM images and absorption spectra shown in Figure 22A and C. The TEM images clearly show large 3D aggregates of particles, and the absorption spectra display a decrease in the dipole resonance at 533

nm associated with an increase in absorbance at longer wavelengths which is consistent with particle aggregation.

Next the SERS kinetic series measured from aggregating particles was considered and the expected trend was observed. The intensity initially increased due to formation of aggregates with strong coupled LSPR modes, followed by the SERS intensity decreasing as the aggregates became too large. An example of how the SERS intensity changes as a function of time is shown in Figure 22D. The rate is estimated by the slope of the Raman kinetic series, as illustrated by the linear fit in Figure 22D. This slope was used as a metric for colloidal stability to assess the effect of CDDP on the nanoparticles. The difference in aggregation rate between blank saline and 0.6 μg/mL CDDP was used to compare the sensitivity of Au-Cys and Au-Cit to the drug. The results shown in Figure 22E show that Au-Cys had a significantly higher responsivity to the drug, and the rate of aggregation was statistically different between treated and untreated particles for both capping agents. Comparing the ζ-potential of the CDDP treated and untreated Au-Cys nanoparticles as shown in Figure 22B, there is a slight shift towards 0 V suggesting that CDDP destabilizes the particles. The mean ζ-potential of the as-synthesized particles was found to be -45 mV, and the peak charge shifts slightly to -43 mV upon interaction with CDDP. Furthermore, there is a significant increase in the charge distribution between -36 and -14 mV in Au-Cys particles with CDDP. At this time, we are unsure if this is caused by CDDP binding to the amine group, or displacing cysteine from the gold nanoparticles by binding to the thiol group. The latter seems more likely based on the observed decrease in surface charge. Based on these results, we have formulated a tentative sensing mechanism illustrated in Figure 23.

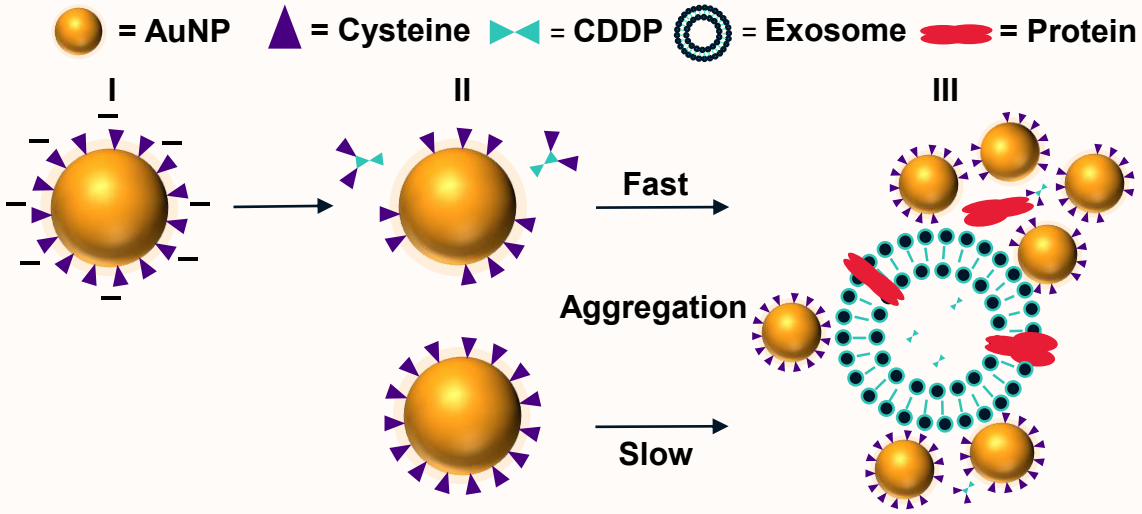

Figure 23: Illustration of sensing mechanism for SERS quantification of CDDP and exosomes. (I) After synthesis, Au-cys particles have a strong negative surface charge which maintains a mono-stable colloid of gold nanoparticles. (II) After incubation with CDDP, the drug binds to the cysteine residues and reduces the surface charge of the particles, thus destabilizing them. In this illustration, it is assumed that the thiol group preferentially conjugates with platinum thus removing the amino acids from the nanoparticle. Aggregation is induced, and the rate at which this happens is proportional to the CDDP concentration. (III) At steady state aggregation, the nanoparticles are stochastically aggregated about exosomes and exosomal proteins/lipids.

Using samples of CDDP in saline as initial proof of concept samples, a regression relating the drug concentration and the aggregation rate was developed. A strong correlation is observed between the particle aggregation rate and CDDP concentration, with the linear range extending from approximately 0 to 0.2 µg/mL as shown in Figure 22F. Further optimization was postponed until exosomes were introduced into the samples in order to account for their effect on the sensor.

5.3.2 Quantifying exosomes and exosome derived CDDP

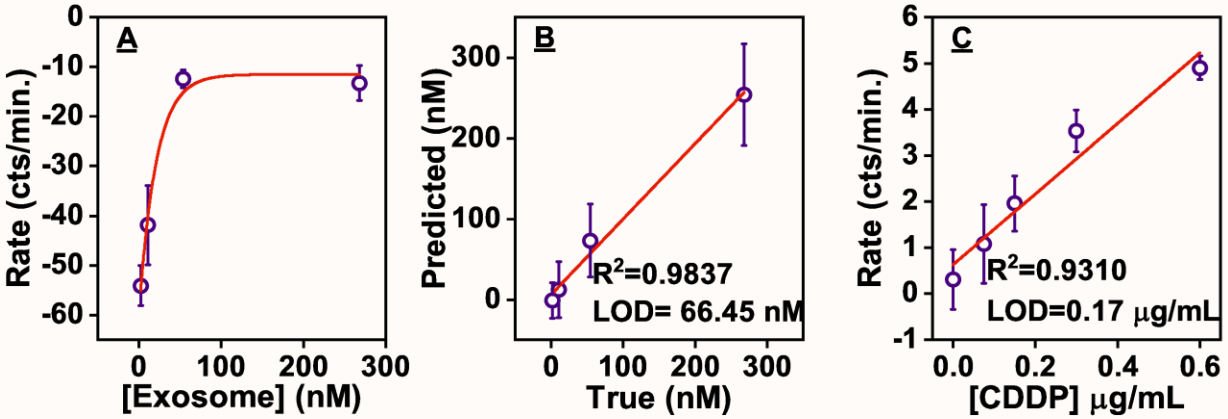

Figure 24: Quantification of exosomes and exosome-derived CDDP. (A) The aggregation rate of the nanoparticles for 0.6 µg/mL CDDP with different concentrations of pooled exosomes. (B) SVR regression model for exosome quantification based on the steady state SERS spectra. (C) Regression relating CDDP concentration to nanoparticle aggregation rate as measured by SERS.

To study the effects of exosomes on the CDDP sensing scheme, initial experiments added pooled samples of exosomes derived from two paired OVCA cell lines, namely A2780S (AS) and A2780CP (CP), which are EM chemosensitive and resistant cell lines respectively. These two cell lines are of special significance because they were originally isolated from the same tumor and only differ in their resistance to chemotherapy. Therefore, we can hypothesize that any differences between these two cell lines are purely a factor of chemoresistance.

First, different concentrations of exosomes were suspended in a 0.6 µg/mL solution of CDDP, and the same aggregation experiments were conduced to observe the effect of the vesicles on the nanoparticles. DLS measurements were used to estimate the initial concentration of the OVCA exosome isolates. To quantify particles with DLS, one requires *a priori* knowledge of the particle refractive index, which for our purposes was taken to be 1.365 based on published data [160]. The concentration of exosomes was found to be 250±40 nM using this method. We found that the exosomes provide a strong stabilizing effect to the Au-Cys particles, even in the presence of CDDP. This is illustrated by the decrease in aggregation rate as the concentration of exosomes increased as shown in Figure 24A. The most probable explanation for this observation is the adsorbed corona of exosome-derived proteins, as this has been shown in other studies to have a stabilizing

effect [161]. Despite the apparent effect of exosomes on our CDDP quantification method, there were two encouraging results. First is that even at high concentrations of exosomes, aggregation was never completely inhibited. Secondly, this effect saturates at exosome concentrations above 100 nM, which means that for exosome samples in the expected range of 250±40 nM variations in the exosome concentration will not have pronounced effects on the aggregation rate. Thus, the stabilizing effects of the exosomes could be accounted for by increasing the amount of NaCl added to cause aggregation, in order to keep the assay in the original 15-minute window.

The formation of a protein corona and the apparent interactions with exosome components suggests that the SERS spectrum should contain vibrational signatures of these molecules. Therefore, the average steady-state SERS spectrum could be used to analyse the concentration and chemical identity of the exosomes using GA-SVM. To this end, a GA-SVR model was built to relate the exosome concentration to the SERS spectrum, and the resultant model is illustrated in Figure 24B. The results indicate that the SVR model is effective in quantifying exosomes with a root mean squared error of 44 nM and a limit of detection of 65 nM. The error of the SVR method is similar to that of our original DLS measurements, indicating that if we were to use a more rigorous quantification method to analyse the exosome standard, then the SERS-based model could also be improved. Future experiments may consider quantifying exosomes using flow cytometry for this reason.

Next, the CDDP detection scheme was recalibrated using undiluted pooled exosome samples from n = 10 independent cell cultures. The quantity of NaCl added was increased from 10 to 50 μL per mL of sample to account for the exosome stabilization effect. The new regression model is shown in Figure 24C, wherein a highly linear relationship (R^2 = 0.93) between the aggregation rate of CDDP concentration was found. The limit of detection for this scheme was found to be 0.17 μg/mL. The most notable difference between this regression and the previous relationship found for CDDP in neat saline is that the aggregation rate is now positive. This is consistent with the observation that the exosomes stabilize the particles, as an earlier stage of the aggregation kinetics was observed in these measurements. Specifically, the exosomes cause the kinetic SERS

measurement to capture monodisperse particles forming optimally sized nanoaggregates, thus resulting in increasing intensity as a function of time. In the earlier results this process occurred too quickly, and instead the growth of large non-resonant aggregates was observed. The CDDP and exosome quantification schemes developed in this section can now be performed simultaneously in a single kinetic SERS measurement to acquire two biomarker concentrations that may relate to chemoresistance.

5.3.3 SVM discriminants to classify exosomal Raman spectra; differentiation between histological subtypes and chemosensitivity

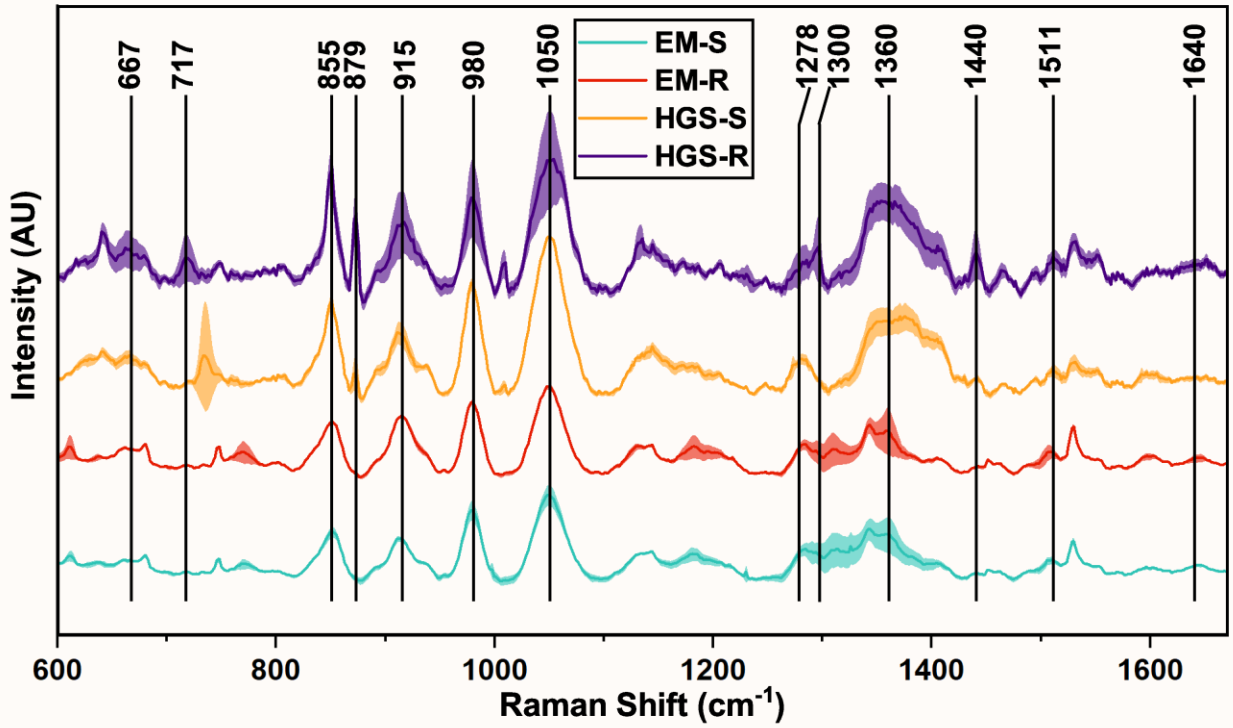

Figure 25: Raman spectra of untreated exosomes on cysteine capped gold nanoparticles acquired from the steady state of the aggregation kinetic series. The shaded regions represent one standard deviation from the mean.

Given the success of SVR based exosome quantification on the Au-Cys SERS substrate, the next test was to find if chemical differences between exosomes could be used to differentiate different histological subtypes and chemoresistance profiles of the associated tumours. This would be possible if there are conserved chemical features

across subtypes, or consistently correlated with chemoresistance. To this end, SERS spectra were acquired from exosomes derived from six OVCA cell lines with different histologic subtypes (HGS and EM) without CDDP treatment. The specific cells used were three EM cell lines, A2780s (AS), A2780cp (CP), and TOV112D (112), as well as three HGS cell lines TOV3041G (3041), TOV3291G (3291), and OV90. Of these, AS and 3041 cells are sensitive to CDDP whereas CP, 112, 3291, and OV90 are resistant. The average exosomal spectra acquired for each of these cell lines using Au-Cys aggregates are shown in Figure 25.

Three SVM discriminant models were then created; one which differentiates between EM and HGS, and two which differentiate between CDDP sensitivity/resistance within each histologic subtype. These results are given in Table 7 and Table 8. It was found that the SVM model could differentiate between EM and HGS ovarian cancer cell derived exosomes with ~75% accuracy. When differentiating between chemosensitive and chemoresistant cell-derived exosomes, the SVM model had significantly different performance between the two histological subtypes. Accuracy was much higher for HGS exosomes (~88%) compared to EM exosomes (~70%), implying that the biochemical differences related to chemosensitivity may be more pronounced in this subtype. However, it should be noted again that the EM set of data contained the paired OVCA cell lines AS and CP. The fact that these are paired cells means that their spectra are more similar, and any differences are more likely to be exclusively correlated to chemoresistance. When comparing exosomes from the other cell lines, spectral differences may be due to heterogeneity between different tumors that is not necessarily correlated to chemoresistance. The accuracy for discriminating between AS and CP exosomes alone also has an accuracy of ~70%, indicating that there are conserved spectral features which differentiate chemoresistance profiles.

Table 7: Confusion matrix showing the prediction of tumor subtype based on exosome SERS spectra

		Predicted ID	
		EM	**HGS**
True ID	**EM**	**75.24**	24.76
	HGS	26.11	**73.89**

Table 8: Confusion matrix showing the prediction of chemoresistance based on exosome SERS spectra separated by subtype

		Predicted (EM/HGS)	
		Sensitive	**Resistant**
True (EM/ HGS)	**Sensitive**	**67.6/85.7**	32.4/14.3
	Resistant	26.7/10.0	**73.3/90.0**

The difference between spectra and first two loading vectors of a partial least squares projection (see Appendix: tentative band assignment for OVCA cell derived exosomes.) were used to determine which Raman bands contributed to differentiating between histologic subtype and chemosensitivity [159,162–165]. Comparing EM and HGS spectra, Raman bands at 1050, 855, and 980 cm^{-1} correlate strongly with HGS. The band at 1050 cm^{-1} likely reflects the H-C-N bending of the cysteine capping molecules, thus indicating differential aggregation behavior of the nanoparticles with respect to subtype specific exosomes. What is more interesting is that the bands at 855 and 980 cm^{-1} are assigned to cytosine/uracil and ribose, respectively. This may suggest that HGS associated exosomes carry more RNA than other subtypes which may be an area for future investigation. With respect to EM cell-derived exosomes, their spectra are more associated with protein bands such as 1505 (Amide II, tryptophan), 1180 (amides), and 769 cm^{-1} (tryptophan). A similar set of bands differentiate chemoresistant and

chemosensitive cell-derived exosome spectra as well. With 1050 and 980 cm^{-1} correlating with chemosensitivity, and 915 (hydroxyproline) and 1306 cm^{-1} (adenine/cytosine ring resonance) correlating with chemoresistance in both HGS and EM. Again, the prevalence of RNA associated bands in these discriminant models points to the possible role of exosome-derived RNAs in ovarian cancer progression, especially in HGS.

5.3.4 Diagnosis of chemoresistance using multiple biomarkers from a single assay

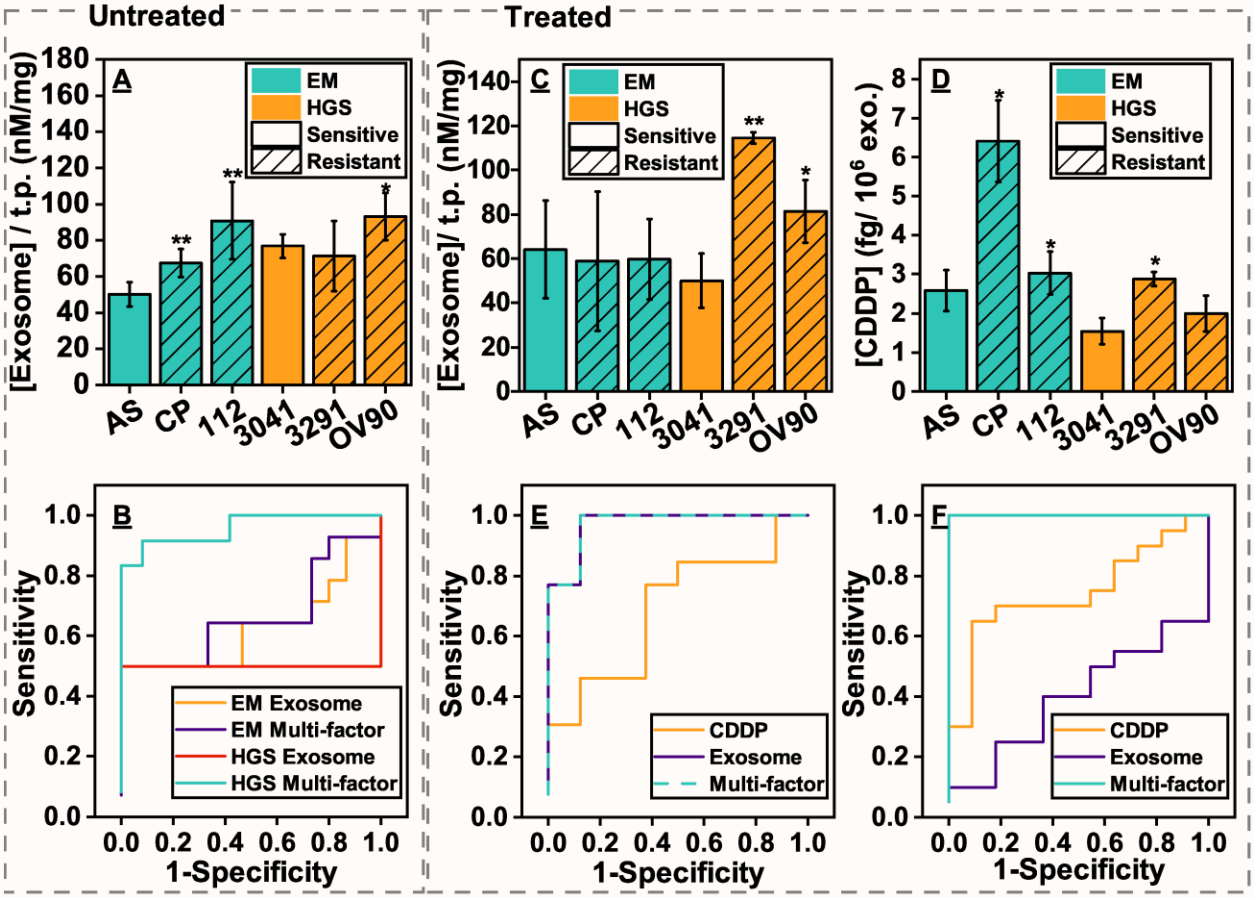

Figure 26: Diagnostic capability of the nanophotonic SERS assay (* is $p \leq 0.05$, ** is $p \leq 10^{-4}$ compared to chemosensitive). (A) Comparing the quantity of exosomes in untreated samples of OVCA cells as measured by SERS. (B) ROC curves for diagnosis of chemoresistance based on quantifying exosomes from untreated EM and HGS cells alone, and multi-factor diagnosis with the SVM prediction. (C) Comparing the quantity of exosomes and (D) CDDP concentration per unit exosome for each OVCA cells treated with CDDP. (E and F) ROC curves for diagnosing chemoresistance based on quantifying exosome, exosomal CDDP, or both combined with SVM prediction for HGS (E) and EM (F).

With the basic characterization of the Au-Cys based kinetic SERS assay complete, we then tested whether a multi-factor diagnostic which combined all previous measurements could act as an effective chemoresistance diagnostic. This included the CDDP concentration from the aggregation rate, the concentration of exosomes from SVR, and the SVM score for chemoresistance discrimination. In order to normalize the measurements, the CDDP measurements were reported as fg/(10^6 exosomes) in order to quantify the rate of sequestration by the OVCA cells. The exosome concentration is divided by the total protein (t.p. in the figures) to account for differences in the number of cells cultured. These metrics were combined with the SVM scores to create a linear multi-factor diagnostic.

Initially, exosomes from untreated cells were analysed to find if the exosome concentration and SVM discriminant could be used as a pre-treatment diagnostic for chemoresistance. The output and receiver operator curve (ROC) from these measurements are shown in Figure 26A and B. The area under the ROC (AUROC) is a metric for the quality of a diagnostic, wherein an AUROC of 0.5 represents a meaningless discriminant and an AUROC of 1 represents a perfect diagnostic. The concentration of exosomes per total protein derived from untreated cells yields a poor diagnostic. Using this metric alone yields AUROCs of 0.63 and 0.5, for EM and HGS respectively. The previous section has shown that SVM discriminant analysis of untreated cells could yield a good diagnostic, particularly for HGS. By combining this with the exosome quantification, the AUROC increases to 0.67 and 0.96 for EM and HGS respectively. This is in agreement with the previous results that indicate SVM is able to discriminate between HGS derived exosomes more readily. From these results we find that diagnosing chemoresistance based on untreated cell-derived exosomes may have some diagnostic power, but the performance is highly subtype dependant.

For the next experiments, CDDP treated samples involved culturing the cells with the drug before isolating exosomes. Therefore, the sensor was quantifying the CDDP sequestered and expelled by the cancer cells during treatment. Upon treatment with CDDP, the exosome quantification results changed significantly as shown in Figure 26C. There was a marked increase in the amount of exosomes released by the chemosensitive

AS cell line, eliminating any difference between chemoresistant and sensitive EM cells. This may be due to the AS strain dying as a result of CDDP treatment and releasing dysfunctional exosomes as debris. Conversely, treatment seems to cause a significant increase in exosomes released by chemoresistant HGS cells. This subtype-dependant difference may be due to mechanistic differences in chemoresistance expression between OVCA subtypes. Instead, if the amount of exosome-derived CDDP is measured, a subtype independent diagnostic was found. It has been previously shown that chemoresistant tumour cells release a higher amount of CDDP in their exosomes than do their sensitive counterparts. The results from CDDP quantification with our kinetic SERS assay tends to agree with this observation. Average CDDP concentrations from chemosensitive EM and HGS derived exosomes are 2.6 and 1.5 fg/(10^6 exosomes) respectively. While for chemoresistant EM and HGS cells, these values increase to 4.7 and 2.5 fg/(10^6 exosomes) respectively. These values are shown in Figure 26D. We found a relatively low statistical significance for this difference ($p \leq 0.05$ for all but OV90) due primarily to the combined variances of the exosome and CDDP quantification. Regardless, we find that CDDP quantification yields a good, subtype independent, diagnostic with AUROCs of 0.75 and 0.69 for EM and HGS respectively. When CDDP and exosome quantification are combined with the SVM scores from the previous section the diagnostic power is improved for both subtypes yielding AUROC of 1 and 0.97 for EM and HGS, respectively.

5.3.5 Studying the effect of plasma gelsolin expression on chemoresistance phenotype by kinetic SERS

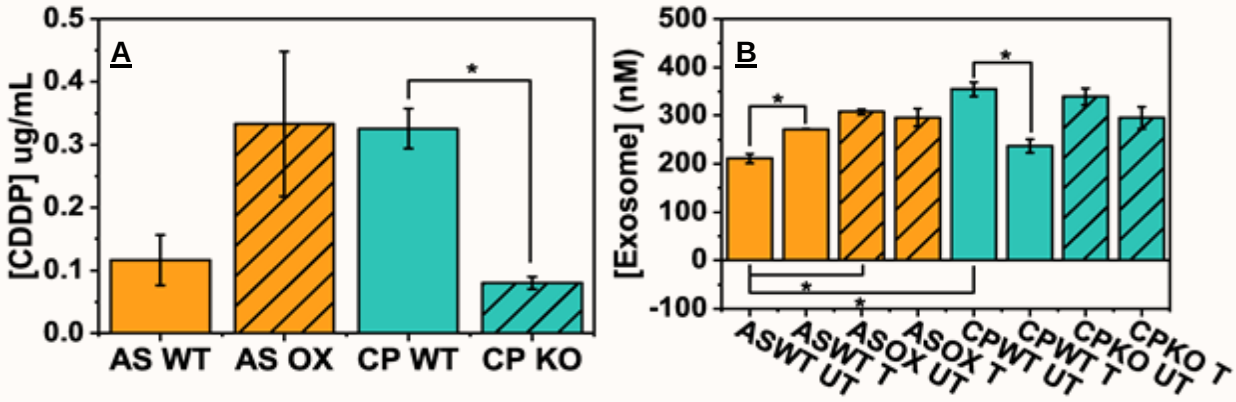

Figure 27: CDDP and exosome quantification from pGSN mutant OVCA cells. (A) CDDP quantification results for chemosensitive AS wild-type (WT), AS pGSN overexpression (OX), chemoresistant CP WT, and CP pGSN knockout (KO). (B) Exosome quantification for the exosomes derived from untreated (UT) and treated (T) OVCA cells. * means $p \leq 0.05$

Previous results from our collaborators at OHRI have identified cell-derived exosomal plasma gelsolin (pGSN) as an important protein in chemoresistance. In healthy individuals, pGSN normally acts as an actin binding protein acting to remove this toxin from the extracellular space. However, Asare-Werehene *et al* have found that pGSN over expression is associated with a shorter time-to-recurrence in OVCA [147]. Knocking down the pGSN gene in chemoresistant cells caused them to become sensitive and undergo apoptosis in response to the drug, and induced overexpression in sensitive cells caused them to become resistant. Furthermore, exosomal secretion of pGSN is increased in chemoresistant cells and these pGSN laden exosomes are able to induce chemoresistance in formerly sensitive cells [147].

Using the chemoresistance diagnostic we had developed in the previous sections, we wished to determine the effect of pGSN expression on exosome production and sequestered CDDP. To this end AS cells were cultured in their natural wild-type (WT) state or with pGSN overexpression (OX), to determine if this induced a resistant phenotype. On the other hand, CT cells were cultured as WT or with pGSN knockdown (KO) in order to eliminate the chemoresistant phenotype. These cells were also either treated with CDDP (T sets) or left untreated (UT sets), in order to find if treatment had an

effect on exosome production in either cell type. The results of CDDP and exosome quantification are shown in Figure 27A and B respectively. For SVM tests, the pGSN mutants were left out of the training set thus making the model naïve to these samples. This was done in order to examine which parent cell, AS or CP, the pGSN mutants were most similar to without biasing the classifier.

The results indicate that overexpression of pGSN in chemosensitive cells (AS) causes increased secretion of both exosomes and exosome derived CDDP. However, this CDDP result is not quite statistically significant due to the large variance of the measurement. pGSN knockout in formerly chemoresistant cells (CP) causes a decrease in CDDP excretion. Therefore, with respect to CDDP sequestration and secretion in the exosomes we find that pGSN plays a role in developing the resistant phenotype. In general, chemoresistant cells tend to excrete more exosomes than do their sensitive counterparts. As expected, when pGSN is overexpressed there is an associated increase in exosome production. However, when pGSN is knocked down in resistant cells there is no decrease in exosome release. Again, this may be due to the CP-KO cells becoming sensitized to CDDP, thus undergoing apoptosis in response to the drug and releasing non-physiological exosomes as debris. With respect to the machine learning results shown in Table 9, all pGSN mutant exosomes are predominantly classified as CP (chemoresistant) using this method. When the cells are treated with CDDP, the discriminant splits slightly between AS and CP, but is still predominantly the latter.

We therefore found that pGSN overexpression induces a resistant phenotype in formerly sensitive cells, but the sensitive phenotype is not fully recovered by pGSN knockdown. Only CDDP excretion levels decrease in response to knockdown, while exosome secretion and overall exosome chemistry do not appear change significantly. The complete mechanisms underlying the development and expression of chemoresistance are still an area of active study. An interesting result from these experiments is that exosome production appears to decrease in the CP-WT cells in response to CDDP. This potentially points to the drug itself as a mechanistic cofactor. There is precedent for this type of interaction, as it has been found that CDDP dissociates cellular gelsolin (cGSN) from an apoptosis inhibiting protein to which it can bind [166].

Table 9: Confusion matrix showing the GA-SVM predicted class for the pGSN mutant OVCA cell-derived exosomes.

		Predicted	
		Sensitive	**Resistant**
True	**CP-pGSN-KO Treated**	80%	20%
	CP-pGSN-KO Untreated	92.5%	7.5%
	AS-pGSN-OX Treated	80%	20%
	AS-pGSN-OX Untreated	100%	0%

5.4 Conclusion

This chapter has described the design of a SERS based diagnostic modality for determining chemoresistance in OVCA. This sensor is based on CDDP modulating the stability of cysteine capped gold nanoparticles, and measuring the inherent SERS spectrum of tumor derived exosomes. CDDP was successfully quantified by measuring the aggregation rate of the nanoparticles based on a kinetic series of Raman spectra. This method achieved quantification of CDDP down to 0.17 μg/mL. The quantity of exosomes could also be acquired in a single measurement by using the exosomal SERS spectrum at the steady state aggregation of the nanoparticles. These spectra were also employed for tumor subtype, and chemosensitivity discrimination based on SVM models. The final result was three diagnostic factors; CDDP concentration, exosome concentration, and SVM class, which could be combined into a multifactor diagnostic. The results show that drug and exosome efflux from tumor cells is both treatment and subtype dependent. In general, we found that chemoresistant OVCA cell lines excrete more exosomes and more exosome-derived CDDP than their sensitive counterparts. However, a single factor was not sufficient to achieve a good diagnostic due to the high variance of tumor cells. The multi-factor diagnostic provided by this SERS sensor was able to overcome this issue, achieving AUROCs of 0.97 to 1. This sensor may be valuable in

monitoring the efficacy of ongoing chemotherapy, and may indicate the development of acquired resistance. While less efficacious, quantifying exosomes and applying SVM discriminant analysis on their Raman spectra prior to CDDP treatment does hold some diagnostic potential. This may be employed as a pre-screening tool in chemotherapy that could be used to avoid ineffective treatment. Finally, we showed the utility of this sensor not only as a potential diagnostic tool, but also as a means of conducting chemoresistance studies by quantifying CDDP and exosomes. This was explored using pGSN mutant OVCA cells, wherein we found that overexpression of this protein leads to the development of the chemoresistant phenotype, and knockdown partially induces the sensitive phenotype.

Chapter 6 Summary and Future Work

6.1 Summary of thesis

The major objectives of this thesis were to advance the field of Raman chemical sensing by exploring three major themes addressed in Chapters 3-5. These were the development of rapid non-linear chemometrics to develop efficacious discriminant and regression models, the design and validation of fiber enhanced SERS flow cytometry as an approach for cellular detection, and the development of a small molecule functionalized SERS substrate to diagnose chemoresistance in ovarian cancers.

Chapter 2 contained the underlying theory and existing knowledge which laid the foundation for the work in this thesis.

Chapter 3 outlined the theory and validation of a support vector machine based system which employed a genetic algorithm to optimized kernel hyperparameters. In this work, a simplistic analytical fitness function was employed that took advantage of the well developed theory underlying SVMs. Additionally, this was the first time that this type of fitness function was applied to support vector regression. The resultant genetic algorithm support vector machine (GA-SVM) was first applied to standard machine learning test sets, and showed comparable results to other reported methods. The time required to optimized the kernel functions was found to be significantly reduced when compared to the conventional grid search method. This algorithm was also faster than other approaches to GA-SVM wherein n-fold cross validation is used as a fitness function, while in this work, the SVM optimization problem only needed to be solved once per model. GA-SVM was found to outperform PLS in all the Raman chemometric regression applications tested. Primarily, the improvements were to the precision of the resultant model, thereby decreasing the total error, with minimal effect on accuracy.

Chapter 4 concerned the development of fiber enhanced SERS flow cytometry with the goal of developing a rapid culture-free methodology for bacterial quantification and identification. This began with some small but significant changes to the HC-PCF flow cell

system developed previously by our group, most notably of which was the change from high performance liquid chromatography manifolds, to those based on bespoke microfluidics. Not only did this halve the system dead volume, but also paves the way for more advance optofluidic systems, some of which being suggested in this chapter. Following basic experiments validating the enhancement and reproducibility afforded by HC-PCF SERS, the system was employed in bacterial detection in fetal bovine serum. The developed method relied on counting spectral events generated by silver decorated bacteria as they flowed through the HC-PCF. Counting these events estimated cellular load and discriminated different species of bacteria using the previously developed GA-SVM. This inherently multiplexed label-free system compared favourably in terms of assay time and detection limit to other methods of bacterial quantification including qPCR and biosensors based on surface plasmon resonance or electrical impedance spectroscopy. This system was also capable of determining the make-up of mixed cultures of bacteria as well, albeit with a decrease in overall performance. We hypothesized that improvements could be made in future iterations of the device by including a viability metric based on the SERS spectra in order to achieve a true measure of the number of colony-forming units in a sample. Finally, methicillin resistant and susceptible *S. aureus* were successfully differentiated in synthetic synovial fluid using the HC-PCF SERS flow cytometer. In these experiments, it was found that a component of the amide II band of the Raman spectrum was elevated in the spectra of susceptible strains, potentially pointing to a difference in how the penicillin binding proteins bind silver nanoparticles. This biological question remains an area for future investigation.

In Chapter 5, a chemo-susceptibility assay was developed for ovarian cancers based on tumor derived exosomes. This technique was based on the cysteine functionalized gold nanoparticles as a SERS substrate. We found that the chemotherapy drug of interest, cisplatin (CDDP), modulated the stability of these particles resulting in a concentration dependant aggregation rate. Cysteine capped particles were found to be much more sensitive to this reaction than citrate, most likely because the former possesses nucleophilic moieties to which CDDP can bind. As the nanoparticles aggregated in the presence of tumor derived exosomes, it was also found that the steady state SERS spectra could be used to quantify these vesicles. Furthermore, using GA-

SVM it was possible to chemically differentiate exosomes originating from different histological subtypes of ovarian cancer and chemoresistance profiles. The result was a three-in-one sensor which could simultaneously acquire information of three biomarkers for chemoresistance in a single assay: exosome derived CDDP concentration, tumor derived exosome concentration, and exosome chemical grouping. These three metrics could be combined as a diagnostic panel to achieve high sensitivity and specificity for chemoresistance diagnosis.

6.2 Novel SERS nanomaterials

6.2.1 Bio-functional metal nanoparticles:

In the work described in Chapter 4, very simple citrate capped silver nanoparticles were employed in SERS detection of bacteria. It seems highly unlikely that citrate would be the best capping agent for the job, and by imparting different bio-functionality we may improve species/strain discrimination of bacteria. Furthermore, different cell detection applications could be developed using different types of nanoparticles that are designed with a specific purpose in mind. There are many families of possible capping agents including small molecules, polymers, and peptides. The cysteine capped gold nanoparticles developed in Chapter 5 would be one example of employing bio-active small molecules as a capping agent to facilitate SERS detection of biological targets. The near infinite scope of possible molecules with which to functionalize a colloidal SERS substrate makes choosing a certain research direction difficult. However, different design philosophies may help narrow down the possibilities. For example, ongoing thesis work by another member of our team is considering whether capping agents that have yielded enhanced bactericidal effects on silver nanoparticles may also improve SERS detection.

Another approach may be to design capping agents using peptides with active motifs to cause desirable biological interactions. Peptide capped nanoparticles are very interesting because a high degree of control over the particle's surface chemistry is obtained by tuning the amino acid sequence. With a well constrained problem, a researcher can develop a peptide sequence that provides desirable bio-stability, cellular uptake, target specific molecules or cells, and modulate the layer of proteins that forms

around nanoparticles in complex biofluids [33]. Peptides for nanoparticle functionalization can fall into one of several families such as stabilizing/bio-inert, cell-penetrating, antimicrobial, receptor, antigen, and enzyme substrate peptides [186]. When a nanoparticle enters biological media, the surface chemistry is immediately affected by the binding of endogenous proteins to the surface, known as the protein corona [33]. In some cases, this may negate desirable properties or biochemical interactions. To avoid the formation of this corona, the capping agent should confer a neutral surface charge, hydrophilicity, and steric hindrance to binding. One common capping molecule that achieves this is polyethylene glycol (PEG), which is very effective at minimizing biological interactions [33]. However, for SERS substrates we do not wish to negate bio-functionality carte-blanche, rather only allow for desired interactions to occur.

Emilio Alarcon's group from the BeATS lab, has developed a peptide motif of sequence cysteine-leucine-lysine (CLK) which possesses a high affinity to metal surfaces [187]. Similar to the single amino acid capping performed in Chapter 5, CLK binds metal nanoparticles primarily though the thiol containing cysteine residue. Addition of arginine-serine (RS) to this sequence results in CLKRS which is hydrophilic and biologically stable much like PEG capped particles [187]. Their group has used CLKRS capped silver nanoparticles as anti-fouling agents in synthetic corneas [188]. We might employ this same metal binding motif to attach bio-functional peptide sequences to the particle surface. For example, there are several cell-penetrating peptides such as poly-arginine (polyR), TAT peptide, and penetratin which have been previously applied to bio-active nanomaterials. Hossain *et al* facilitated binding of TAT and penetratin to gold nanoparticles by adding a terminal cysteine residue [189]. Their goal was to create a theranostic SERS system by additionally capping the particles with PEG for stability and the drug doxorubicin. Their design relied on these cell penetrating peptides to cause the drug laden particles to enter cancer cells, and then the diffusion of the drug into the cells could be monitored by SERS. During these experiments, they found that TAT was more effective at introducing the particles into cells than was penetratin. As another example, Shi *et al* used TAT functionalized gold nanostars as a SERS probe to study the biochemical changes which occur as mesenchymal stem cells differentiate[190]. To my knowledge there have been no efforts to employ polyR capped nanoparticles in SERS

thus far, however Tanvir *et al* showed that polyR/PVP capped silver nanoprisms possessed powerful bactericidal properties [191]. Due to the fact that most cells are negatively charged, polyR facilitates cell penetration in part by giving the particles a positive surface charge. There are also anti-microbial peptides that may be employed in a SERS sensor for bacteria. For example, Yuan *et al* developed magnetic nanoparticles capped with the peptide bacitracin that could isolate bacteria from whole blood [192]. Subsequent mixing of the isolated bacteria with gold/silver graphene oxide nanoparticles allowed for intrinsic SERS detection down to 10 CFU/mL with 97.3% discrimination accuracy. Another class of peptides is known as “antigen motif” sequences, which mimic natural ligands to target proteins of interest. These have been employed extensively in other biosensors [186], but have yet to be used in SERS to my knowledge.

The peptide sequence CLK-polyR, could be used to improve bacteria detection of HC-PCF SERS flow cytometry by facilitating penetration of the SERS substrate into the cells. For these experiments, gold should be used in place of silver to avoid cellular toxicity effects. There are many variables involved in synthesizing and using these particles that would need to be optimized to produce an improved sensor. These variables include the number of arginine residues in the polyR chain, the ratio of peptide to additional stabilizer (e.g. PEG-SH, PVP, or CLKRS), and the concentrations of capping molecules, reducing agent, and gold ions. Additionally, it may be valuable to initially pursue several possible synthesis routes such as gold nanoparticle formation by citrate, ascorbic acid, sodium borohydride, or a photo-initiator. The experimental challenge would be best addressed by employing a partial-factorial experiment design in order to minimize the number of experiments needed. The interested reader is encouraged to consult the excellent text by Selvamuthu and Das regarding design of experiments for a foundational understanding of this approach [193]. In the case of bacteria identification, the desired output we seek to maximize is the discriminant ability of the classifier. This can be quantified by both the resultant accuracy, and the kernelized Mahalanobis distance. The latter metric defines how separated different classes are in the high dimensional feature space. Therefore, if the distance is high then classes are well separated and thus easily distinguishable, the opposite being true in the case of a low distance. Following this optimization, the physiochemical properties of the final colloid should be analysed. Morphology can be

determined by TEM imaging, the SPR wavelength by UV-Vis spectroscopy, and the zeta potential by dynamic light scattering.

6.2.2 Core-shell magnetic particles with molecularly imprinted polymers

Molecularly imprinted polymers (MIPs) are synthetic bio-recognition elements that have been gaining significant attention. Much like antibodies or aptamers, MIPs possess binding domains with high specificity towards a desired antigen. The binding regions are cavities that are formed in the polymer during synthesis due to the presence of the target molecule in the reaction. The template molecules are then removed using solvents to create empty cavities that are ready to be used as recognition elements. MIPS are quite inexpensive and relatively easy to fabricate, but they have lower target specificity than antibodies or aptamers. These materials can take a variety of forms including thin films of polymer, and polymer microspheres. However, I believe that the form most useful to our research is core-shell magnetic nanoparticle MIPs (MMIPs). This is because they allow for enrichment of the target molecule by isolating the MMIPs from solution using a magnetic field. The result is a significantly concentrated target molecule with which we can leverage SERS chemometrics to offset the confounding cross-reactivity of MIPs and specifically quantify the target. Additionally, target molecules can easily be removed from the MIP binding sites, thus allowing for cleaning and reuse of the SERS substrate which could significantly reduce the amount of consumables used in a given assay.

In recent years, there have been a number of reports on combining MIPs and SERS for small molecule sensor development. Ashley *et al* synthesized MMIPs using the antibiotic cloxacillin as a template molecule, and then bound these polymers to magnetic nanoparticles [29]. These particles were added to porcine serum and incubated to allow cloxacillin to bind to the MIPs. The particles were then removed from the serum by magnetic separation, then dispersed in a solution of methanol/acetic acid which caused cloxacillin to be removed from the MIPs. This solution was then dried onto a SERS substrate consisting of a nano-pillar array etched into an Si substrate and coated with gold and the Raman spectrum was acquired. This study is an example of a two-stage technique, wherein MMIPs are initially employed to isolate the target before it is released and detected by SERS. There are also examples of MIPs and SERS substrates being

combined onto a single particle. Hu and Lu created a melamine sensor using a MIP cavity that contained both a melamine binding site as well as several AgNPs [194]. After removal of the template molecule, the MIP-AgNPs could capture melamine and provide a strong SERS signal of the molecule. Zhou *et al* developed sensor for phenylalanine which employed silver nanoparticles coated with MIP films [24]. The MIP film on the silver nanoparticles ensured that phenylalanine molecules were close to the surface of the particles and thus within the LSPR field. This system achieved recovery rates of greater than 95% for detecting phenylalanine in serum. It should be noted that all of these aforementioned applications involve MIPs created for small molecule detection. Biomacromolecules such as proteins present a challenge to MIP synthesis due to their relative size [195]. At this time, there have been significant advancements in synthesizing core-shell magnetic MIPs for biomacromolecules, but there have yet to be any SERS sensors based on these particles to the my knowledge [196].

MMIP based sensors could be employed in our current research in projects where the detection/quantification of a single target is required. It is likely that intrinsic, label-free, approaches supported by GA-SVM will continue to yield good results for differentiating between chemically distinct classes of biological structures. However, this technique may struggle to quantify a distinct target in a complex matrix. One example application would be to quantify pGSN in ovarian cancer patients. In Chapter 5, the effects of pGSN expression on chemoresistance in ovarian cancer cell lines was investigated using the CDDP/exosome sensor developed in that chapter. However, our collaborators on this project have also discovered that quantifying pGSN itself could be a biomarker of chemoresistance [147]. Additionally, their group is also interested in discovering novel diagnostic modalities for poly-cystic ovarian syndrome (PCOS). This disease occurs in up to 10% of females but is significantly under-diagnosed. Suffering from untreated PCOS significantly increases the likelihood of infertility, diabetes, and cancer [197]. Previously, our former colleague Dr. Momenpour developed a SERS scheme to detect chemerin as a diagnostic marker for PCOS [111]. Dr. Tsang's group has also identified neuropeptide

Y as a potential marker for PCOS that may be used in conjunction, or in place of, chemerin [198].

A good starting point regarding the synthesis of gold-coated MMIPs is the procedure developed by Li *et al*, which can be modified slightly to create additional nanoparticles after initial MIP fabrication [199]. A tentative synthesis is illustrated in Figure 28. Initially, iron ions (Fe^{2+} and Fe^{3+}) are reduced by sodium citrate in a water/ammonia mixture to form magnetic nanoparticles (I). Again using citrate as a reducing agent, gold ions can be reduced onto the iron surfaces (II). After magnetically isolating the particles, they can be

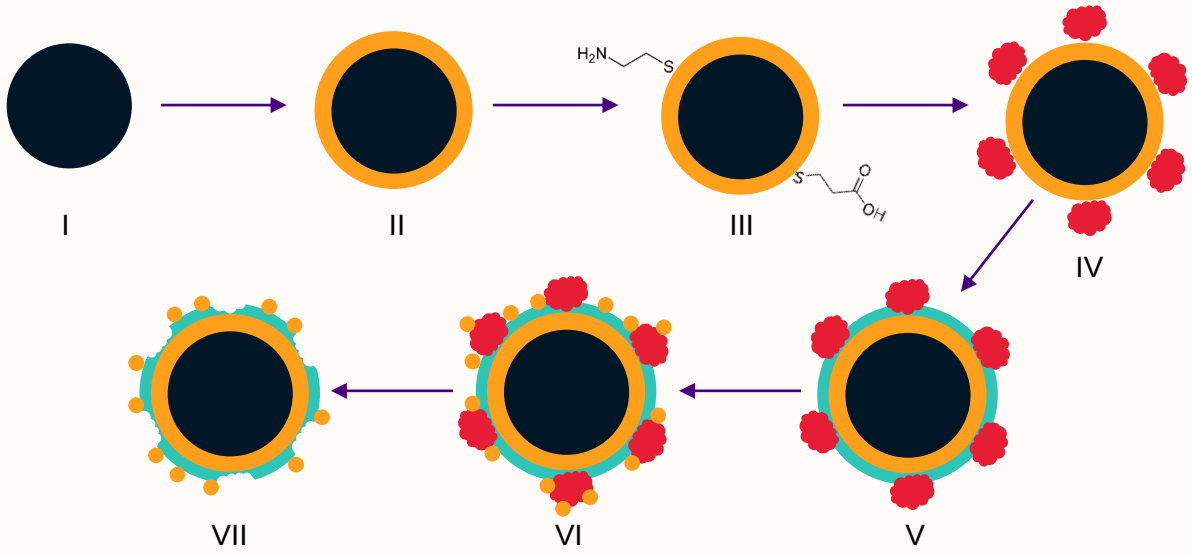

Figure 28: Possible MIP synthesis route. (I) Magnetic iron nanoparticles (MNPs) are synthesized. (II) MNPs are coated with gold. (III) Gold surfaces are functionalized with different binding moieties. (IV) Target analyte is introduced and allowed to bind to particles. (V) Poly-dopamine layer is formed on the particle surface to form the MIP. (VI) Additional gold nanoparticles are reduced onto the MIP layer. (VII) Target is removed to produce sensing element.

suspended in an HCl solution. This will dissolve any bare iron particles, while the less reactive gold will etch much slower. The particles can now by cleaned by magnetically isolating them and re-suspending in clean DIW several times. Next, active carboxyl and amine moieties are required at the particle surface in order to bind the target protein. This can be achieved by adding molecules such as mercaptopropionic acid or cysteamine to the nanoparticle solution (III). The target can then be added and allowed to bind to the particle surfaces (IV). One can then apply dopamine as the basis for the MIP layer, as it will auto-polymerize on the particle surfaces (V). In order to create a SERS substrate

capable of resolving intrinsic protein SERS spectra, additional nanoparticles can be reduced onto the MIP surface (VI). The template protein can then be removed by washing the particles with a solution of methanol and acetic acid (VII). Optimization of this synthesis is a multivariate problem concerning the concentrations of iron, gold, sodium citrate, template protein, dopamine, thiols, as well as the reaction time and temperature. This can be achieved using an appropriate DOE method, as briefly discussed in the previous section. To start, R6G can be used instead of proteins to develop a functional synthesis before concerning oneself with the challenges of making macromolecule MIPs. By using MMIPs in conjunction with the HC-PCF flow cell, one should be able to significantly enrich a target protein of interest and assay a large fluid volume in order to create a sensor with a very low detection limit. This will likely be necessary for detecting trace protein markers in blood. Additionally, GA-SVR can be applied in order to compensate for non-linear effects arising from non-specific binding to the MIPs.

6.3 Continued development of machine learning chemometrics

The GA-SVM system developed in Chapter 3 served the work conducted in this thesis quite well. However, there are several ways in which our spectral processing algorithms could be improved. While discriminant and regression models were generated via SVM in this thesis, spectral band assignment and feature visualization was performed using PLS. This is because the kernel function, while computationally expedient, obscures the data representation in the feature space. For this reason, it is non-trivial to visualize the importance of input features to the SVM model. It would be valuable to be able to do the type of analysis shown in Section 5.3.3 for example using SVM, because this would ideally consider non-linearly weighted features. Additionally, the removal of unnecessary features would reduce the final size of the model, making it easier to implement on integrated logic/microcontrollers in an in-line sensing system. Wrapper methods are one class of feature extraction algorithms more amenable to non-linear data [200]. Wrapper methods use a specific machine learning algorithm and learned model to generate feature weights by iteratively trying all combinations of possible features. There are many examples of this type of algorithm, including the popular recursive feature elimination (RFE) method which can used with SVMs to select features from non-linear models [201].

SVM-RFE has been successfully applied to Raman spectra by Kampe *et al* where they showed that RFE could successfully reduce the dimensionality of spectral data without compromising model accuracy [202]. It is possible that by using RFE weighted spectra, we may visualize the importance of difference Raman bands in SVM models. This would result in the ability to perform band assignment analysis using this non-linear method.

SVMs are also supervised learning machines which require operator labelled datasets to train a given model. There are many cases however where manually labelled spectra may be difficult or impossible to acquire. For example, consider automated pixel classicisation in a Raman image, where we may wish to false-colour the picture based on chemical composition. Applying a supervised approach would require the operator to manually assign a class to each pixel (ex. collagen, cell nucleus, cell membrane, etc.) to train a model, which would introduce bias and error if the operator misinterpreted an image. Instead, it may be valuable to apply an unsupervised approach wherein the data labels are learned by an automatic clustering algorithm. There are several notable examples of these methods being used in Raman spectral analysis including work by Taylor *et al* who used unsupervised ML methods to discriminate between healthy and cancerous cells using sub-cellular Raman images [203]. Recently, an algorithm called unsupervised feature selection extreme machine learning (UFSEML) has been proposed which combines both clustering and feature selection into a single algorithm [204]. The results from Chen *et al* suggest that this method outperforms previously developed unsupervised methods when applied to standard test data used in ML development [204].

In addition to developing new and improved spectral processing methods, it will also be valuable to integrate them with existing open-source spectral analysis packages. This will improve the usability, and the reach of the algorithms we have developed and foster broad collaborations. One recently developed processing library for MATLAB may be a good place to start [205].

6.4 SERS microscopy

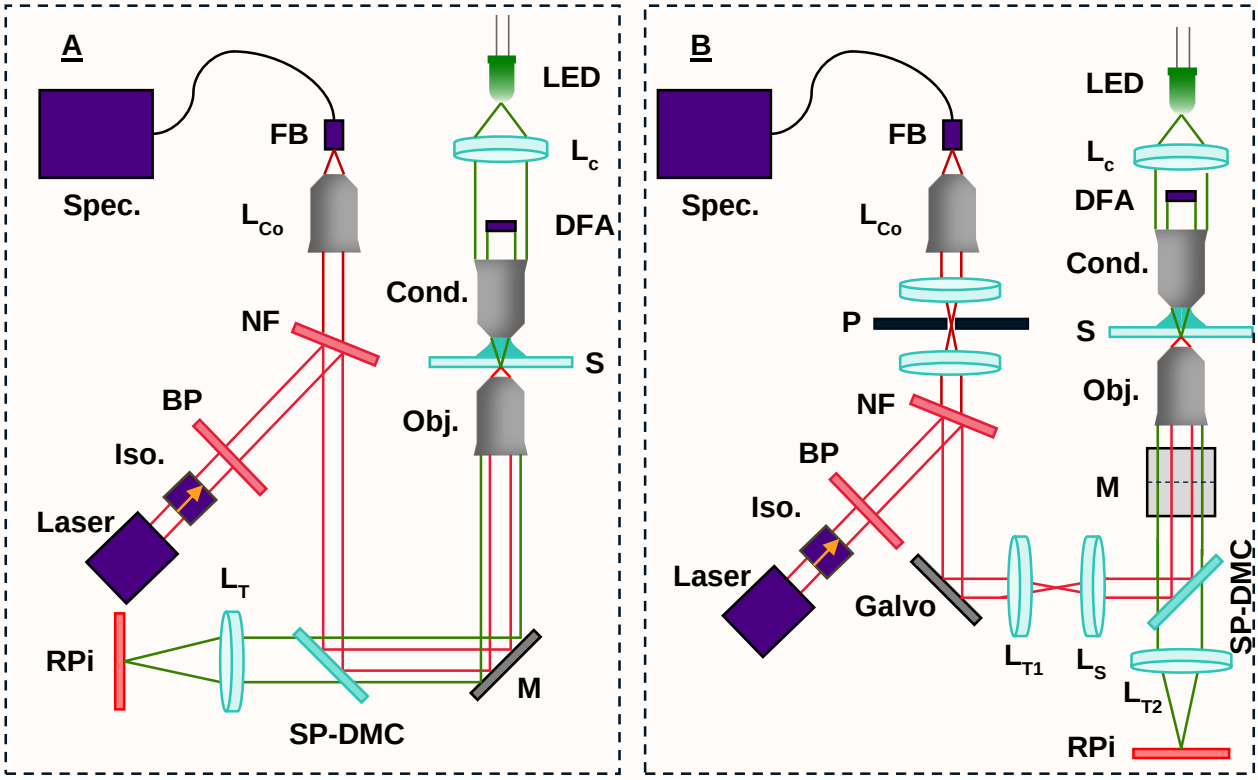

Figure 29: Schematics for basic Raman microscopy setups. (A) Widefield (bright or dark) microscope with integrated Raman spectrometer. (B) Scanning laser microscope with combined widefield system

Research conducted in parallel to the projects discussed in this thesis concerned the development of complimentary SERS microscopy techniques to continue the development of our existing technologies and support new research directions and applications. The first simple design was a combined bright/dark field microscope with a Raman spectrometer in order to perform highly targeted single cell spectroscopy. The schematic of this setup is shown in Figure 29, wherein the microscope portion consists of LED illumination with a collector lens (L_C), that is focused by the condenser onto the sample. The condenser may also be fitted with a darkfield stop as pictured here. For darkfield microscopy, the numerical aperture of the condenser must be greater than that of the objective, and as such a water submerged condenser is employed here (NA = 0.85). The image of the sample is collected by an objective lens (NA = 0.65), which passes through a shortpass dichroic mirror in the microscope's infinity space before the image is formed on a Raspberry Pi (RPi) camera via the tube lens (L_T). The RPi runs custom

Python code that controls the camera, and the images are visualized on an integrated touchscreen alongside the user interface. The Raman spectrometer is similar to that used in other work in this thesis, with one notable difference being the notch filter (NF) that replaces the usual longpass dichroic mirror. This device's stopband is centred at 808 nm when the incident wavefront is parallel to the filter. However, the 785 nm laser employed here is positioned 22° with respect to the filter, thus causing it to efficiently reflect the excitation wavelength. This microscope is operated with green illumination due to the high scattering cross-section of gold nanoparticles illuminated at their resonant wavelength. This facilitates viewing of nanoparticle labelled structures very easily in darkfield.

Data collected from this device will be the subject of a future thesis from our team, so no further experimental details will be discussed here. Single cell spectroscopy on fixed specimens compliments the development of the SERS flow cytometry system because it facilitates facile measurements without needing to be concerned with actively flowing cells. However, this spectrometer integrated microscope is also very limited, as it can only acquire spectra from a single spot in the image plane. For applications involving larger biological structures such as eukaryotic cells or whole tissue, it would be valuable to construct full SERS images. For example, results from Taylor *et al* suggest that using sub-cellular Raman spectra of cells can improve the ability to discriminate between subtypes [203]. In order to add full SERS microscopy to the current device, the most straightforward route would be to construct a standard confocal scanning system. This involves the addition of a galvo-scanning mirror assembly and a pair of scan lenses to the system, as well as timing control over the mirrors and the spectrometer. These changes are illustrated in Figure 29B, wherein the microscope portion is built perpendicular to the optical bench, thus all optics from the objective to the LED should be visualized as coming out of the page.

While the ability to acquire SERS images using a confocal scanning system would be valuable for a number of research avenues, the system itself is far from novel. The use of confocal Raman microscopy is well established in academic practice. A more interesting approach, which should also yield superior resolution and faster imaging times, is to instead implement a structured illumination microscopy (SIM) system. This

approach is a widefield technique, which means that the image of the entire sample area is acquired at once rather than by scanning a laser beam waist and constructing the image pixel by pixel. The result is that image acquisition times can be significantly reduced using this technique. Widefield acquisition of Raman spectral images means that using a conventional spectrometer is no longer feasible. Rather than spatially separating spectral components using a grating, individual frequencies are imaged one at a time using tunable optical filters. Thus, in the case of widefield Raman hyperspectral imaging, spatial images are generated in a single acquisition while spectral data is acquired serially, whereas the opposite is true for confocal scanning Raman microscopy.

Widefield Raman images could be captured by illuminating the sample with a flat intensity distribution laser beam and then collecting Raman scattered photons. However, in this approach the resolution of the microscope is diffraction limited, and the axial sectioning capabilities of confocal scanning are lost, meaning that out-of-focus scattered light will blur the image. This is where structured illumination enters the picture. Rather than using an isotropic light distribution, samples are illuminated with a coherent harmonic pattern resulting in low frequency Moiré fringes from the high spatial frequency components of the image [206,207]. By capturing at least one image per spatial frequency component of the illumination pattern at different phases, it is possible to mathematically reconstruct the high spatial frequencies of the sample from the Moiré fringes, resulting in up to a two-fold increase in spatial resolution. In these systems, axial resolution is achieved by fixing the illumination pattern with respect to the focal plane of the objective [208]. To generate these structured illumination patterns, a spatial light modulator (SLM) [209], digital micro-mirror device (DMD) [210,211], or phase grating [208] can be employed. There have been several realizations of Raman SIM in recent years such as the device developed by Chen *et al.* which used an SLM and a set of three notch filters to acquire Raman SIM images of labelled nanoparticles [212]. Wang *et al* designed a Raman SIM system using a DMD to generate the spatial patterns, along with tunable long and shortpass filters which were used in tandem to create an effective narrow bandpass filter [213]. This microscope achieved a spectral resolution of 50 cm^{-1} and a spatial resolution of 80 nm.

To illustrate the potential impact of developing a Raman SIM system for our research, consider the Basler acA2000-165umNIR camera we have in the lab which has a resolution of two megapixels. Also consider our ongoing research regarding SERS analysis of tissue, wherein the integration time required to get a spectrum with good SNR is one minute. In this case, to get a single Raman image using a scanning confocal system with equivalent resolution to the Basler camera would take around four years. This is probably impractical, and clinicians and patients alike would likely be opposed to multi-year diagnostic turn-around times. Instead let us imagine a Raman SIM system which uses conventional 3-beam illumination, resulting in five independent spatial frequency components. If the system had a spectral resolution of 1 cm^{-1} over the full fingerprint region (~100 – 1800 cm^{-1}), the acquisition of a single image would require a few days to complete assuming a similar integration time. This could be further reduced to a clinically reasonable time by selecting a smaller subset of Raman bands to image, optimizing the SERS substrate, and/or employing immuno-SERS techniques [214].

Appendix: tentative band assignment for OVCA cell derived exosomes.

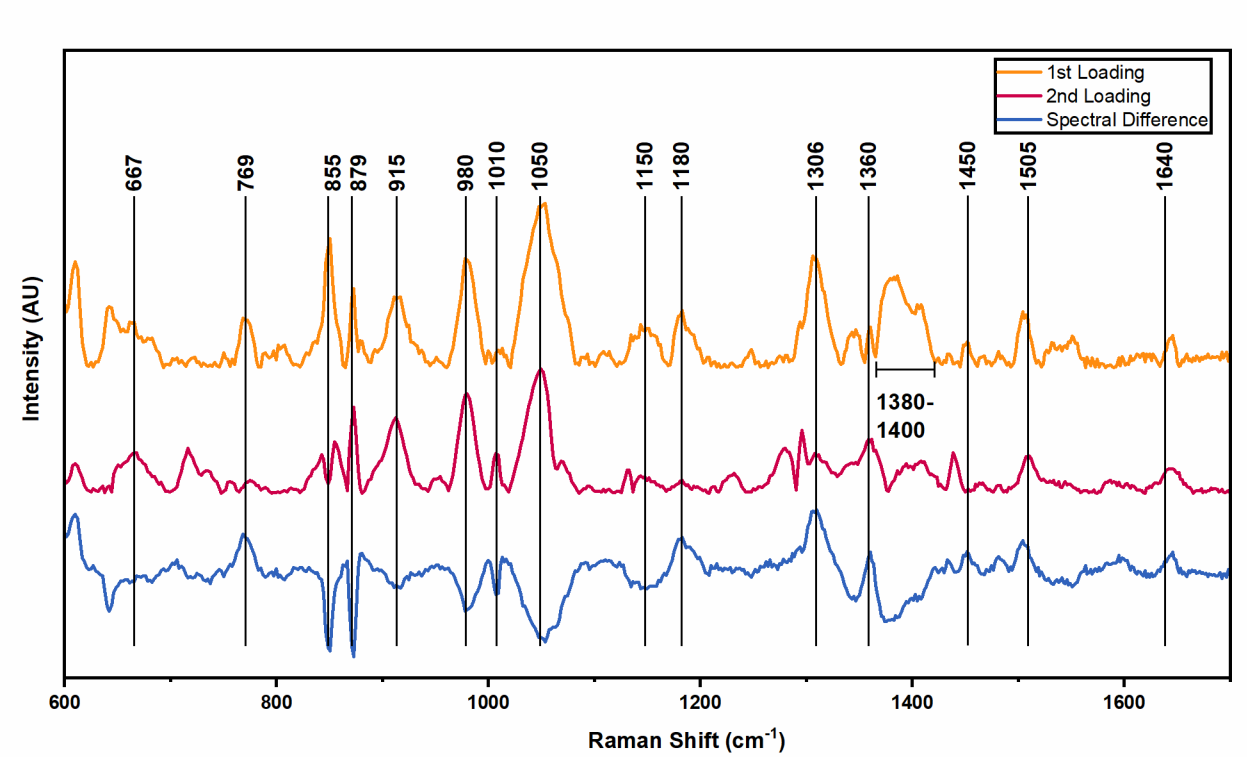

Figure 30: PLS loading vectors and difference spectra (EM – HGS) describing the difference between the two subtypes of OVCA explored in these experiments; high-grade serous and endometrioid.

Table 10: List of important Raman bands, ranked by weight in the first loading vector, differentiating EM and HGS spectra.

Raman Band (cm^{-1})	Weight (Loading 1)	Weight (Loading 2)	Assignment[159,162–165]	Association
1050	2.20492	1.63829	C-N, amides, Phenylalanine, DNA, adenine	HGS

855	1.74125	0.18643	Tyrosine, Cytosine, Uracil	HGS
1306	1.5097	0.48878	C-N, N-H, adenine, DNA	EM
980	1.47249	1.34601	Ribose	HGS
1380-1400	1.24554	0.30651	CH_3	HGS
879	1.06806	1.16577	C-C, hydroxyproline	HGS
915	0.95739	0.84474	$C\text{-}COO^-$ stretch	HGS
1505	0.71341	0.49695	Amide II, Tryptophan	EM
1180	0.70569	0.16553	C-O stretch, amides, protein	EM
769	0.67101	0.07616	Tryptophan ring breathing	EM
1360	0.54475	0.69212	Adenine ring	EM
1150	0.53321	0.20674	C-N stretch, amides, DNA, adenine	HGS
667	0.45439	0.54771	Nucleic Acids, C-S stretch (Cysteine)	Marginal
1640	0.39399	0.32682	Amide I	EM

1450	0.34775	0.04112	CH_2 deformation, lipids, proteins, carbohydrates	EM
1010	0.23833	0.50752	Phenylalanine	HGS

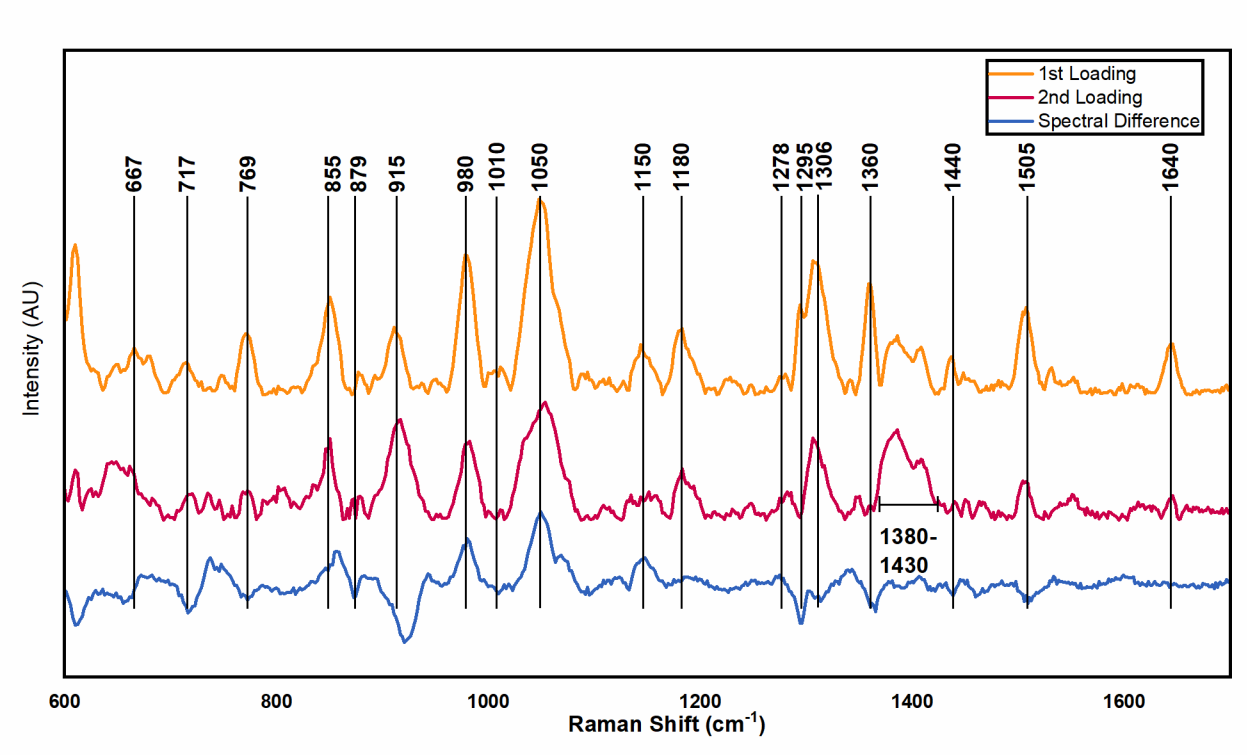

Figure 31: PLS loading vectors and difference spectra (susceptible – resistant) describing the spectral difference between chemoresistant and susceptible OVCA cell derived exosomes from endometrioid cells.

Table 11: List of important Raman bands, ranked by weight in the first loading vector, differentiating the spectra of exosomes derived from susceptible and resistant endometrioid cell.

Raman Band (cm^{-1})	Weight (Loading 1)	Weight (Loading 2)	Assignment[159,162–165]	Association

1050	1.86806	1.04141	C-N stretch, amides, Phenylalanine, DNA, adenine	Susceptible
980	1.32967	0.72638	Ribose	Susceptible
1306	1.28283	0.78228	C-N, N-H, adenine, DNA	Resistant
1360	1.05474	0.14248	Adenine ring	Resistant
855	0.83996	0.51337	Tyrosine, Cytosine, Uracil	Susceptible
1505	0.83445	0.36998	Amide II, Tryptophan	Resistant
1180	0.63061	0.49317	C-O stretch, amides, protein	Susceptible
915	0.60365	0.90992	C-COO- stretch	Resistant
769	0.5845	0.25187	Tryptophan ring breathing	Resistant
1380-1430	0.52507	0.8325	CH3	Susceptible
1640	0.48381	0.1999	Amide I	Marginal
1150	0.47865	0.22168	C-N stretch, amides, DNA, adenine	Susceptible
667	0.39906	0.30909	Nucleic Acids, C-S stretch (Cysteine)	Resistant
1440	0.36498	0.10627	C-H (Lipids), CH2 deformation, lipids, proteins, carbohydrates	Resistant
717	0.30804	0.21368	C-S (Cysteine), Tryptophan	Resistant

1010	0.26576	0.09697	Phenylalanine	Marginal
879	0.21535	0.2188	C-C, hydroxyproline	Resistant
1278	0.19649	0.24109	Amide III	Susceptible

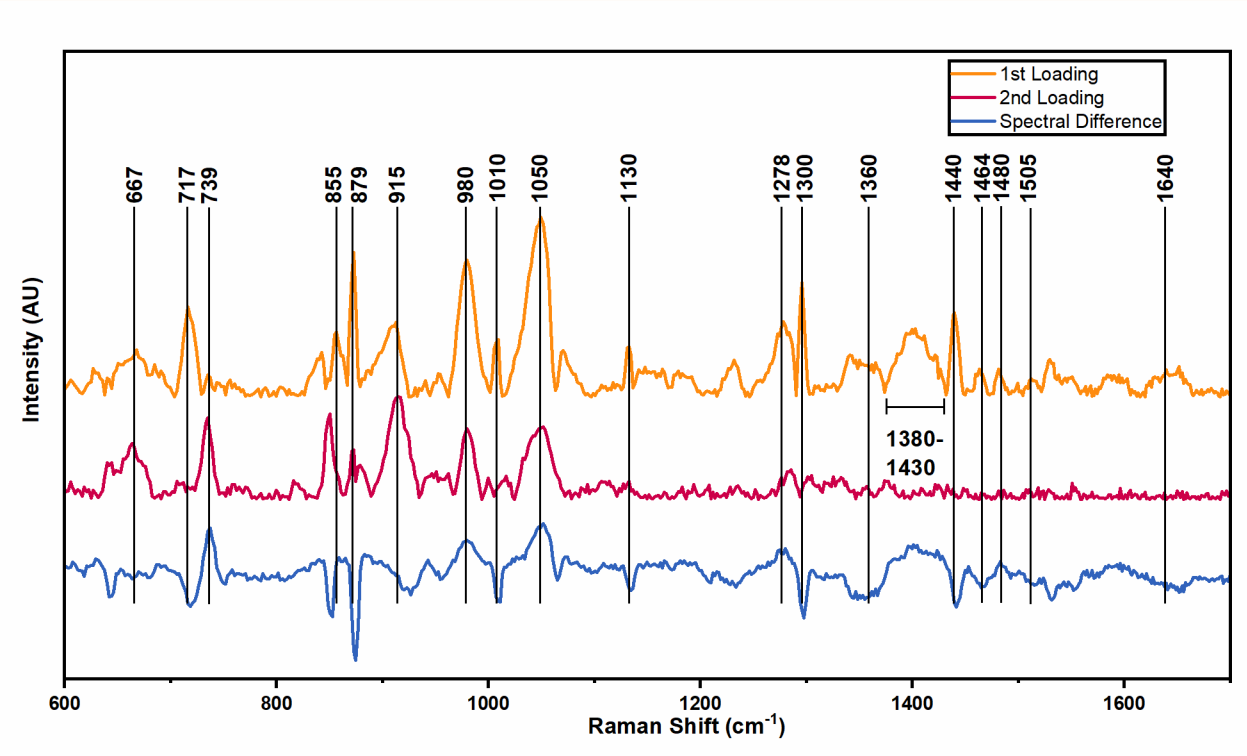

Figure 32: PLS loading vectors and difference spectra (susceptible – resistant) describing the spectral difference between chemoresistant and susceptible OVCA cell derived exosomes from high-grade serous cells.

Table 12: List of important Raman bands, ranked by weight in the first loading vector, differentiating the spectra of exosomes derived from susceptible and resistant high-grade serous cell.

Raman Band (cm^{-1})	Weight (Loading 1)	Weight (Loading 2)	Assignment[159,162–165]	Association

1050	1.4024	0.55746	C-N stretch, amides, Phenylalanine, DNA, adenine	Susceptible
915	1.12595	0.38486	C-C, hydroxyproline	Resistant
980	1.06625	0.55322	Ribose	Susceptible
1306	0.8911	0.0761	C-N, N-H, adenine, DNA	Resistant
717	0.70375	0.09651	C-S (Cysteine), Tryptophan	Resistant
1440	0.65863	0.02139	C-H (Lipids), CH2 deformation, lipids, proteins, carbohydrates	Resistant
1278	0.58409	0.11345	Amide III	Sensitive
915	0.55399	0.76051	C-COO- stretch	Resistant
1380-1430	0.53322	0.04567	CH3	Susceptible
855	0.50686	0.21049	Tyrosine, Cytosine, Uracil	Resistant
1010	0.44396	0.093	Phenylalanine	Sensitive
1130	0.38077	0.12111	C-N, Proline	Resistant
667	0.36853	0.33216	Nucleic Acids, C-S stretch (Cysteine)	Marginal
1360	0.32706	0.05791	Adenine ring	Resistant
1464	0.21804	0	C-H (Lipids)	Resistant

1480	0.21488	0.06308	Histidine	Susceptible
1640	0.2129	0	Amide I	Marginal
739	0.17236	0.63645	C-S (Cysteine), Tryptophan	Susceptible
1505	0.14185	0.05319	Amide II, Tryptophan	Resistant

References

[1] B. Schrader, Infrared and Raman Spectroscopy: Methods and Applications, 2008. doi:10.1524/zpch.1998.205.part_1.127a.

[2] P.J. Larkin, IR and Raman Spectroscopy: Principals and Spectral Interpretation, Elsevier Inc., 2011.

[3] G. Turrell, J. Corset, Raman Microscopy, 1996. doi:10.1002/9783527619283.ch3a.

[4] M. Okuno, Hyper-Raman spectroscopy of alcohols excited at 532 nm: Methanol, ethanol, 1-propanol, and 2-propanol, J. Raman Spectrosc. 52 (2021) 849–856. doi:10.1002/jrs.6066.

[5] A. Taghizadeh, U. Leffers, T.G. Pedersen, K.S. Thygesen, A library of ab initio Raman spectra for automated identification of 2D materials, Nat. Commun. 11 (2020) 1–10. doi:10.1038/s41467-020-16529-6.

[6] J. Carroll, J. Whiteaway, D. Plumb, Distributed Feedback Semiconductor Lasers, 1998. doi:10.11316/butsuri1946.31.265.

[7] D. Lin, S. Qiu, W. Huang, J. Pan, Z. Xu, R. Chen, S. Feng, G. Chen, Y. Li, M. Short, J. Zhao, Y. Fawzy, H. Zeng, Autofluorescence and white light imaging-guided endoscopic Raman and diffuse reflectance spectroscopy for in vivo nasopharyngeal cancer detection, J. Biophotonics. 11 (2018) 1–9. doi:10.1002/jbio.201700251.

[8] S. Mosca, C. Conti, N. Stone, P. Matousek, Spatially offset Raman spectroscopy, Nat. Rev. Methods Prim. 1 (2021). doi:10.1038/s43586-021-00019-0.

[9] Q.Y. Li, K. Takahashi, X. Zhang, Frequency-domain Raman method to measure thermal diffusivity of one-dimensional microfibers and nanowires, Int. J. Heat Mass Transf. 134 (2019) 539–546. doi:10.1016/j.ijheatmasstransfer.2019.01.057.

[10] J.W. Chan, Recent advances in laser tweezers Raman spectroscopy (LTRS) for label-free analysis of single cells, J. Biophotonics. 6 (2013) 36–48. doi:10.1002/jbio.201200143.

[11] A. Shamsaie, M. Jonczyk, J. Sturgis, J. Paul Robinson, J. Irudayaraj, Intracellularly grown gold nanoparticles as potential surface-enhanced Raman scattering probes, J. Biomed. Opt. 12 (2007) 020502. doi:10.1117/1.2717549.

[12] P. Hermann, A. Hermelink, V. Lausch, G. Holland, M. Lars, D. Naumann, Evaluation of tip-enhanced Raman spectroscopy for characterizing different virus strains †, (2011) 1148–1152. doi:10.1039/c0an00531b.

[13] N.E. Markina, S.N. Ustinov, A.M. Zakharevich, A. V. Markin, Copper nanoparticles for SERS-based determination of some cephalosporin antibiotics in spiked human urine, Anal. Chim. Acta. 1138 (2020) 9–17. doi:10.1016/j.aca.2020.09.016.

[14] G. Sun, J.B. Khurgin, Plasmonic Enhancement of Optical Properties By Isolated and Coupled Metal Nanoparticles, in: Plasmon. Plasmonic Metamaterials Anal. Appl., 2011: pp. 1–44. doi:10.1142/9789814355285_0001.

[15] S. V. Gaponenko, Introduction to nanophotonics, 2010. doi:10.1017/CBO9780511750502.

[16] D.W. Shipp, F. Sinjab, I. Notingher, Raman spectroscopy: techniques and applications in the life sciences, Adv. Opt. Photonics. 9 (2017) 315. doi:10.1364/aop.9.000315.

[17] K. Tanabe, Field enhancement around metal nanoparticles and nanoshells: A systematic investigation, J. Phys. Chem. C. 112 (2008) 15721–15728. doi:10.1021/jp8060009.

[18] E.W. Weisstein, Laplace’s Equation--Spherical Coordinates., MathWorld - A Wolfram Web Resour. (n.d.). https://mathworld.wolfram.com/LaplacesEquationSphericalCoordinates.html.

[19] N. Engheta, A. Salandrino, A. Alù, Circuit elements at optical frequencies: Nanoinductors, nanocapacitors, and nanoresistors, Phys. Rev. Lett. 95 (2005) 1–4. doi:10.1103/PhysRevLett.95.095504.

[20] E.M. Purcell, Spontaneous emission probabilities at radiofrequencies, Phys. Rev. 69 (1946) 681. http://link.aps.org/doi/10.1103/PhysRev.69.674.

[21] J. Moon, S.Y. Yi, A. Hwang, G. Eom, J. Sim, J. Jeong, E.K. Lim, B.H. Chung, B. Kim, J. Jung, T. Kang, Facile and sensitive detection of influenza viruses using SERS antibody probes, RSC Adv. 6 (2016) 84415–84419. doi:10.1039/c6ra13966c.

[22] S. Díaz-Amaya, L.K. Lin, A.J. Deering, L.A. Stanciu, Aptamer-based SERS biosensor for whole cell analytical detection of E. coli O157:H7, Anal. Chim. Acta. 1081 (2019) 146–156. doi:10.1016/j.aca.2019.07.028.

[23] Y. Pang, C. Wang, L.C. Lu, C. Wang, Z. Sun, R. Xiao, Dual-SERS biosensor for one-step detection of microRNAs in exosome and residual plasma of blood samples for diagnosing pancreatic cancer, Biosens. Bioelectron. 130 (2019) 204–213. doi:10.1016/j.bios.2019.01.039.

[24] J. Zhou, S. Sheth, H. Zhou, Q. Song, Highly selective detection of L-Phenylalanine by molecularly imprinted polymers coated Au nanoparticles via surface-enhanced Raman scattering, Talanta. 211 (2020) 120745. doi:10.1016/j.talanta.2020.120745.

[25] S. Schlücker, B. Küstner, A. Punge, R. Bonfig, A. Marx, P. Ströbel, Immuno-Raman microspectroscopy: In situ detection of antigens in tissue specimens by surface-enhanced Raman scattering, J. Raman Spectrosc. 37 (2006) 719–721. doi:10.1002/jrs.1534.

[26] X. Fu, Z. Cheng, J. Yu, P. Choo, L. Chen, J. Choo, A SERS-based lateral flow assay biosensor for highly sensitive detection of HIV-1 DNA, Biosens. Bioelectron. 78 (2016) 530–537. doi:10.1016/j.bios.2015.11.099.

[27] Z. Li, Y. Yi, X. Luo, N. Xiong, Y. Liu, S. Li, R. Sun, Y. Wang, B. Hu, W. Chen, Y.

Zhang, J. Wang, B. Huang, Y. Lin, J. Yang, W. Cai, X. Wang, J. Cheng, Z. Chen, K. Sun, W. Pan, Z. Zhan, L. Chen, F. Ye, Development and Clinical Application of A Rapid IgM-IgG Combined Antibody Test for SARS-CoV-2 Infection Diagnosis., J. Med. Virol. (2020) 0–1. doi:10.1002/jmv.25727.

[28] Y. Wang, Y. Xia, Bottom-up and top-down approaches to the synthesis of monodispersed spherical colloids of low melting-point metals, Nano Lett. 4 (2004) 2047–2050. doi:10.1021/nl048689j.

[29] J. Ashley, K. Wu, M.F. Hansen, M.S. Schmidt, A. Boisen, Y. Sun, Quantitative detection of trace level cloxacillin in food samples using magnetic molecularly imprinted polymer extraction and surface-Enhanced raman spectroscopy nanopillars, Anal. Chem. 89 (2017) 11484–11490. doi:10.1021/acs.analchem.7b02725.

[30] I. Aleknavičienė, E. Pabrėža, M. Talaikis, M. Jankunec, G. Račiukaitis, Low-cost SERS substrate featuring laser-ablated amorphous nanostructure, Appl. Surf. Sci. 571 (2022). doi:10.1016/j.apsusc.2021.151248.

[31] J. Polte, Fundamental growth principles of colloidal metal nanoparticles - a new perspective, CrystEngComm. 17 (2015) 6809–6830. doi:10.1039/c5ce01014d.

[32] M. Wuithschick, A. Birnbaum, S. Witte, M. Sztucki, U. Vainio, N. Pinna, K. Rademann, F. Emmerling, R. Kraehnert, J. Polte, Turkevich in New Robes: Key Questions Answered for the Most Common Gold Nanoparticle Synthesis, ACS Nano. 9 (2015) 7052–7071. doi:10.1021/acsnano.5b01579.

[33] C.D. Spicer, C. Jumeaux, B. Gupta, M.M. Stevens, Peptide and protein nanoparticle conjugates: Versatile platforms for biomedical applications, Chem. Soc. Rev. 47 (2018) 3574–3620. doi:10.1039/c7cs00877e.

[34] M. Ahumada, C. Lazurko, E.I. Alarcon, Fundamental concepts on surface chemistry of nanomaterials, Elsevier Inc., 2019. doi:10.1016/B978-0-12-814531-9.00001-4.

[35] Q. Zhang, N. Li, J. Goebl, Z. Lu, Y. Yin, A systematic study of the synthesis of silver

nanoplates: Is citrate a “magic” reagent?, J. Am. Chem. Soc. 133 (2011) 18931–18939. doi:10.1021/ja2080345.

[36] J.E. Park, Y. Lee, J.M. Nam, Precisely Shaped, Uniformly Formed Gold Nanocubes with Ultrahigh Reproducibility in Single-Particle Scattering and Surface-Enhanced Raman Scattering, Nano Lett. 18 (2018) 6475–6482. doi:10.1021/acs.nanolett.8b02973.

[37] P. Li, Y. Wu, D. Li, X. Su, C. Luo, Y. Wang, J. Hu, G. Li, H. Jiang, W. Zhang, Seed-Mediated Synthesis of Tunable-Aspect-Ratio Gold Nanorods for Near-Infrared Photoacoustic Imaging, Nanoscale Res. Lett. 13 (2018) 0–7. doi:10.1186/s11671-018-2734-8.

[38] F. Liebig, R. Henning, R.M. Sarhan, C. Prietzel, C.N.Z. Schmitt, M. Bargheer, J. Koetz, A simple one-step procedure to synthesise gold nanostars in concentrated aqueous surfactant solutions, RSC Adv. 9 (2019) 23633–23641. doi:10.1039/c9ra02384d.

[39] Q.K. Vo, M.N.N. Thi, P.P.N. Thi, D.T. Nguyen, Green and facile synthesis of dendritic and branched gold nanoparticles by Gelatin and investigation of their biocompatibility on fibroblast cells, Processes. 7 (2019). doi:10.3390/pr7090631.

[40] A. Samanta, S. Jana, R.K. Das, Y.T. Chang, Wavelength and shape dependent SERS study to develop ultrasensitive nanotags for imaging of cancer cells, RSC Adv. 4 (2014) 12415–12421. doi:10.1039/c3ra46208k.

[41] S. Bhaskar, R. Patra, N.C.S.S. Kowshik, K.M. Ganesh, V. Srinivasan, P. Chandran S, S.S. Ramamurthy, Nanostructure effect on quenching and dequenching of quantum emitters on surface plasmon-coupled interface: A comparative analysis using gold nanospheres and nanostars, Phys. E Low-Dimensional Syst. Nanostructures. 124 (2020) 114276. doi:10.1016/j.physe.2020.114276.

[42] Y.Y. Cai, J.G. Liu, L.J. Tauzin, D. Huang, E. Sung, H. Zhang, A. Joplin, W.S. Chang, P. Nordlander, S. Link, Photoluminescence of Gold Nanorods: Purcell Effect

Enhanced Emission from Hot Carriers, ACS Nano. 12 (2018) 976–985. doi:10.1021/acsnano.7b07402.

[43] F. Benabid, J.C. Knight, G. Antonopoulos, P.S.J. Russell, Stimulated Raman Scattering in Hydrogen-Filled Hollow-Core Photonic Crystal Fiber, Science (80-.). 298 (2002) 399–402.

[44] P.S.J. Russell, Photonic-Crystal Fibers, J. Light. Technol. 24 (2006) 4729–4749.

[45] A. Khetani, J. Riordon, V. Tiwari, A. Momenpour, M. Godin, H. Anis, Hollow core photonic crystal fiber as a reusable Raman biosensor, Opt. Express. 21 (2013) 367–370. doi:10.1364/OE.21.012340.

[46] G. Antonopoulos, F. Benabid, T.A. Birks, D.M. Bird, J.C. Knight, P.S.J. Russell, Experimental demonstration of the frequency shift of bandgaps in photonic crystal fibers due to refractive index scaling, Opt. Express. 14 (2006) 3000. doi:10.1364/oe.14.003000.

[47] K. Nielsen, D. Noordegraaf, T. Sørensen, A. Bjarklev, T.P. Hansen, Selective filling of photonic crystal fibres, J. Opt. A Pure Appl. Opt. 7 (2005) L13–L20. doi:10.1088/1464-4258/7/8/l02.

[48] Y. Huang, Y. Xu, A. Yariv, Fabrication of functional microstructured optical fibers through a selective-filling technique, Appl. Phys. Lett. 85 (2004) 5182–5184. doi:10.1063/1.1828593.

[49] Y. Zhang, C. Shi, C. Gu, L. Seballos, J.Z. Zhang, Liquid core photonic crystal fiber sensor based on surface enhanced Raman scattering, Appl. Phys. Lett. 90 (2007) 2005–2008. doi:10.1063/1.2738185.

[50] D. Yan, J. Popp, M.W. Pletz, T. Frosch, Highly sensitive broadband Raman sensing of antibiotics in step-index hollow-core photonic crystal fibers, ACS Photonics. 4 (2017) 138–145. doi:10.1021/acsphotonics.6b00688.

[51] J. Dybas, T. Chiura, K.M. Marzec, P.J. Mak, Probing Heme Active Sites of

Hemoglobin in Functional Red Blood Cells Using Resonance Raman Spectroscopy, J. Phys. Chem. B. 125 (2021) 3556–3565. doi:10.1021/acs.jpcb.1c01199.

[52] R.A. Karaballi, A. Nel, S. Krishnan, J. Blackburn, C.L. Brosseau, Development of an electrochemical surface-enhanced Raman spectroscopy (EC-SERS) aptasensor for direct detection of DNA hybridization, Phys. Chem. Chem. Phys. 17 (2015) 21356–21363. doi:10.1039/c4cp05077k.

[53] R.W. Boyd, Nonlinear Optics, 2003. doi:10.1016/S0074-6142(99)80001-1.

[54] A.E. Kandjani, M.J. Griffin, R. Ramanathan, S.J. Ippolito, S.K. Bhargava, V. Bansal, A new paradigm for signal processing of Raman spectra using a smoothing free algorithm : Coupling continuous wavelet transform with signal removal method, J. Raman Spectrosc. 44 (2013) 608–621. doi:10.1002/jrs.4232.

[55] M.A. Arnold, G.W. Small, Noninvasive glucose sensing, Anal. Chem. 77 (2005) 5429–5439. doi:10.1021/ac050429e.

[56] P. Geladi, Notes on the history and nature of partial least squares (PLS) modelling, J. Chemom. 2 (1988) 231–246. doi:10.1002/cem.1180020403.

[57] M. Barker, W. Rayens, Partial least squares for discrimination, J. Chemom. 17 (2003) 166–173. doi:10.1002/cem.785.

[58] L.A. Bratchenko, I.A. Bratchenko, Y.A. Khristoforova, D.N. Artemyev, D.Y. Konovalova, P.A. Lebedev, V.P. Zakharov, Raman spectroscopy of human skin for kidney failure detection, J. Biophotonics. 14 (2021) 1–10. doi:10.1002/jbio.202000360.

[59] V. Vapnik, The Nature of Statistical Learning Theory, 1995.

[60] M. Saleem, S. Ali, M.B. Khan, A. Amin, M. Bilal, H. Nawaz, M. Hassan, Optical diagnosis of hepatitis B virus infection in blood plasma using Raman spectroscopy and chemometric techniques, J. Raman Spectrosc. 51 (2020) 1067–1077.

doi:10.1002/jrs.5896.

[61] T. Sciortino, R. Secoli, E. D'amico, S. Moccia, M.C. Nibali, L. Gay, M. Rossi, N. Pecco, A. Castellano, E. De Momi, B. Fernandes, M. Riva, L. Bello, Raman spectroscopy and machine learning for idh genotyping of unprocessed glioma biopsies, Cancers (Basel). 13 (2021) 1–13. doi:10.3390/cancers13164196.

[62] E. Ryzhikova, N.M. Ralbovsky, V. Sikirzhytski, O. Kazakov, L. Halamkova, J. Quinn, E.A. Zimmerman, I.K. Lednev, Raman spectroscopy and machine learning for biomedical applications: Alzheimer's disease diagnosis based on the analysis of cerebrospinal fluid, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 248 (2021) 119188. doi:10.1016/j.saa.2020.119188.

[63] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, A Practical Guide to Support Vector Classification, BJU Int. 101 (2008) 1396–1400. http://www.csie.ntu.edu.tw/%7B~%7Dcjlin/papers/guide/guide.pdf.

[64] A. Khetani, V.S. Tiwari, A. Harb, H. Anis, Monitoring of heparin concentration in serum by Raman spectroscopy within hollow core photonic crystal fiber, Opt. Express. 19 (2011) 15244. doi:10.1364/oe.19.015244.

[65] A.M.T. Monfared, V.S. Tiwari, V.L. Trudeau, H. Anis, Surface-enhanced raman scattering spectroscopy for the detection of glutamate and γ-Aminobutyric acid in serum by partial least squares analysis, IEEE Photonics J. 7 (2015). doi:10.1109/JPHOT.2015.2423284.

[66] A. Khetani, A. Momenpour, E.I. Alarcon, H. Anis, Hollow core photonic crystal fiber for monitoring leukemia cells using surface enhanced Raman scattering (SERS), Opt. Express. 23 (2015) 4599–4609. doi:10.1364/OE.23.04599.

[67] Sijmen Jong, SIMPLS : an alternative approach to partial least squares regression, Chemom. Intell. Lab. Syst. 18 (1993) 251–263.

[68] J. Shawe-Taylor, N. Christianini, Kernel Methods for Pattern Analysis, 2004.

[69] A. Neumaier, Solving ill-conditioned and singular linear systems: A tutorial on regularization, SIAM Rev. 40 (1998) 636–666. doi:10.1137/S0036144597321909.

[70] A.J. Smola, B. Schölkopf, A Tutorial on Support Vector Regression, Stat. Comput. 14 (2004) 199–222. doi:Doi 10.1023/B:Stco.0000035301.49549.88.

[71] T.P. Runarsson, Asynchronous Parallel Evolutionary Model Selection for Support Vector Machines, Processing. 3 (2004) 59–68.

[72] J. Platt, Sequential minimal optimization: A fast algorithm for training support vector machines. Advances in Kernel Methods-Support Vector learning, Cambridge, MA MIT Press. (1999) 185–208.

[73] J.N. Hwang, S.R. Lay, M. Maechler, R.D. Martin, J. Schimert, Regression Modeling in Back-propagation and Projection Pursuit Learning, IEEE Trans. Neural Networks. 5 (1994) 342–353. doi:10.1109/72.286906.

[74] T. Baglin, T.W. Barrowcliffe, A. Cohen, M. Greaves, Guidelines on the use and monitoring of heparin, Br. J. Haematol. 133 (2006) 19–34. doi:10.1111/j.1365-2141.2005.05953.x.

[75] C. Advokat, A.I. Pellegrin, Excitatory amino acids and memory: Evidence from research on Alzheimer's disease and behavioral pharmacology, Neurosci. Biobehav. Rev. 16 (1992) 13–24. doi:10.1016/S0149-7634(05)80046-6.

[76] D.A. Greenburg, Glutamate and Parkinson's Disease, Ann. Neurol. 35 (1994) 639.

[77] H. Dohner, D.J. Weisdorf, C.D. Bloomfield, Acute Myeloid Leukemia, N. Engl. J. Med. 372 (2015) 1136–1152. doi:10.1002/9781118010136.ch5.

[78] B. Mohlenhoff, M. Romeo, M. Diem, B.R. Wood, Mie-type scattering and non-Beer-Lambert absorption behavior of human cells in infrared microspectroscopy, Biophys. J. 88 (2005) 3635–3640. doi:10.1529/biophysj.104.057950.

[79] World Health Organization, Report on the Burden of Endemic Health Care-

Associated Infection Worldwide, 2011.

[80] A. Zaucha-Prażmo, J.R. Kowalczyk, K. Drabko, K. Czyżewski, J. Goździk, O. Zając-Spychała, J. Wachowiak, J. Frączkiewicz, E. Gorczyńska, K. Kałwak, J. Styczyński, Incidence of Infectious Complications in Children With Acute Lymphoblastic Leukemia Treated With Hematopoietic Stem Cell Transplantation, Transplant. Proc. 49 (2017) 2183–2187. doi:10.1016/j.transproceed.2017.09.027.

[81] S.B. Whittle, K.C. Williamson, H. V. Russell, Incidence and risk factors of bacterial and fungal infection during induction chemotherapy for high-risk neuroblastoma, Pediatr. Hematol. Oncol. 34 (2017) 331–342. doi:10.1080/08880018.2017.1396386.

[82] A.W. Friedrich, Control of hospital acquired infections and antimicrobial resistance in Europe : the way to go, Wiener Medizinische Wochenschrift. 169 (2019) 25–30. doi:10.1007/s10354-018-0676-5.

[83] C.F. Lowe, K. Katz, A.J. Mcgeer, M.P. Muller, Toronto ESBL Working Group, Efficacy of Admission Screening for Extended-Spectrum Beta-Lactamase Producing Enterobacteriaceae, PLoS One. 8 (2013) 1–8. doi:10.1371/journal.pone.0062678.

[84] A. Trampuz, K.E. Piper, M.J. Jacobson, A.D. Hanssen, K.K. Unni, D.R. Osmon, J.N. Mandrekar, F.R. Cockerill, J.M. Steckelberg, J.F. Greenleaf, R. Patel, Sonication of Removed Hip and Knee Prostheses for Diagnosis of Infection, N. Engl. J. Med. 357 (2007) 654–663. doi:10.1056/nejmoa061588.

[85] A.J. Tande, R. Patel, Prosthetic joint infection, Clin. Microbiol. Rev. 27 (2014) 302–345. doi:10.1128/CMR.00111-13.

[86] W. Gu, S. Miller, C.Y. Chiu, Clinical Metagenomic Next-Generation Sequencing for Pathogen Detection, Annu. Rev. Pathol. Mech. Dis. 14 (2019) 319–338. doi:10.1146/annurev-pathmechdis-012418-012751.

[87] M. Cižman, T. Plankar Srovin, Antibiotic consumption and resistance of gram-

negative pathogens (collateral damage)., GMS Infect. Dis. 6 (2018) Doc05. doi:10.3205/id000040.

[88] M.G. Quiles, L.C. Menezes, K. de C. Bauab, E.K. Gumpl, T.T. Rocchetti, F.S. Palomo, F. Carlesse, A.C.C. Pignatari, Diagnosis of bacteremia in pediatric oncologic patients by in-house real-time PCR, BMC Infect. Dis. 15 (2015) 1–8. doi:10.1186/s12879-015-1033-6.

[89] K. Becker, A.R. Larsen, R.L. Skov, G.K. Paterson, M.A. Holmes, A.J. Sabat, A.W. Friedrich, R. Kočk, G. Peters, A. Kriegeskorte, Evaluation of a modular multiplex-PCR methicillin-resistant staphylococcus aureus detection assay adapted for mecC detection, J. Clin. Microbiol. 51 (2013) 1917–1919. doi:10.1128/JCM.00075-13.

[90] J. Gomez-Cruz, S. Nair, A. Manjarrez-Hernandez, S. Gavilanes-Parra, G. Ascanio, C. Escobedo, Cost-effective flow-through nanohole array-based biosensing platform for the label-free detection of uropathogenic E. coli in real time, Biosens. Bioelectron. 106 (2018) 105–110. doi:10.1016/j.bios.2018.01.055.

[91] H. Vaisocherová-Lísalová, I. Víšová, M.L. Ermini, T. Špringer, X.C. Song, J. Mrázek, J. Lamačová, N. Scott Lynn, P. Šedivák, J. Homola, Low-fouling surface plasmon resonance biosensor for multi-step detection of foodborne bacterial pathogens in complex food samples, Biosens. Bioelectron. 80 (2016) 84–90. doi:10.1016/j.bios.2016.01.040.

[92] S.M. Yoo, D.K. Kim, S.Y. Lee, Aptamer-functionalized localized surface plasmon resonance sensor for the multiplexed detection of different bacterial species, Talanta. 132 (2015) 112–127. doi:10.1016/j.talanta.2014.09.003.

[93] I. Yazgan, N.M. Noah, O. Toure, S. Zhang, O.A. Sadik, Biosensor for selective detection of E. coli in spinach using the strong affinity of derivatized mannose with fimbrial lectin, Biosens. Bioelectron. 61 (2014) 266–273. doi:10.1016/j.bios.2014.05.008.

[94] K. Bekir, H. Barhoumi, M. Braiek, A. Chrouda, N. Zine, N. Abid, A. Maaref, A.

Bakhrouf, H. Ben Ouada, N. Jaffrezic-Renault, H. Ben Mansour, Electrochemical impedance immunosensor for rapid detection of stressed pathogenic Staphylococcus aureus bacteria, Environ. Sci. Pollut. Res. 22 (2015) 15796–15803. doi:10.1007/s11356-015-4761-7.

[95] N. Pal, S. Sharma, S. Gupta, Sensitive and rapid detection of pathogenic bacteria in small volumes using impedance spectroscopy technique, Biosens. Bioelectron. 77 (2016) 270–276. doi:10.1016/j.bios.2015.09.037.

[96] H. Yang, H. Zhou, H. Hao, Q. Gong, K. Nie, Detection of Escherichia coli with a label-free impedimetric biosensor based on lectin functionalized mixed self-assembled monolayer, Sensors Actuators, B Chem. 229 (2016) 297–304. doi:10.1016/j.snb.2015.08.034.

[97] L. Sarcina, E. Macchia, G. Loconsole, G. D'Attoma, P. Saldarelli, V. Elicio, G. Palazzo, L. Torsi, Surface Plasmon Resonance Assay for Label-Free and Selective Detection of Xylella Fastidiosa , Adv. NanoBiomed Res. 1 (2021) 2100043. doi:10.1002/anbr.202100043.

[98] D. Mondal, R. Binish, S. Samanta, D. Paul, S. Mukherji, Detection of Total Bacterial Load in Water Samples Using a Disposable Impedimetric Sensor, IEEE Sens. J. 20 (2020) 1712–1720. doi:10.1109/JSEN.2019.2950422.

[99] S.P. Ravindranath, Y. Wang, J. Irudayaraj, SERS driven cross-platform based multiplex pathogen detection, Sensors Actuators B. Chem. 152 (2011) 183–190. doi:10.1016/j.snb.2010.12.005.

[100] H. Zhang, X. Ma, Y. Liu, N. Duan, S. Wu, Z. Wang, B. Xu, Gold nanoparticles enhanced SERS aptasensor for the simultaneous detection of Salmonella typhimurium and Staphylococcus aureus, Biosens. Bioelectron. 74 (2015) 872–877. doi:10.1016/j.bios.2015.07.033.

[101] N.A. Mungroo, G. Oliveira, S. Neethirajan, SERS based point-of-care detection of food-borne pathogens, Microchim. Acta. 183 (2015) 697–707. doi:10.1007/s00604-

015-1698-y.

[102] A. Sivanesan, E. Witkowska, W. Adamkiewicz, Ł. Dziewit, A. Kamińska, J. Waluk, Nanostructured silver-gold bimetallic SERS substrates for selective identification of bacteria in human blood, Analyst. 139 (2014) 1037–1043. doi:10.1039/c3an01924a.

[103] C. Wang, F. Madiyar, C. Yu, J. Li, Detection of extremely low concentration waterborne pathogen using a multiplexing self-referencing SERS microfluidic biosensor, J. Biol. Eng. 11 (2017) 1–11. doi:10.1186/s13036-017-0051-x.

[104] J.E.L. Villa, N.R. Quiñones, F. Fantinatti-Garboggini, R.J. Poppi, Fast discrimination of bacteria using a filter paper–based SERS platform and PLS-DA with uncertainty estimation, Anal. Bioanal. Chem. 411 (2019) 705–713. doi:10.1007/s00216-018-1485-9.

[105] C.M. Maclaughlin, N. Mullaithilaga, G. Yang, S.Y. Ip, C. Wang, G.C. Walker, Surface-Enhanced Raman Scattering Dye-Labeled Au Nanoparticles for Triplexed Detection of Leukemia and Lymphoma Cells and SERS Flow Cytometry, Langmuir. 29 (2013) 1908−1919. doi:10.1021/la303931c.

[106] G. Perozziello, P. Candeloro, A. De Grazia, F. Esposito, M. Allione, M.L. Coluccio, R. Tallerico, I. Valpapuram, L. Tirinato, G. Das, A. Giugni, B. Torre, P. Veltri, U. Kruhne, G. Della Valle, E. Di Fabrizio, Microfluidic device for continuous single cells analysis via Raman spectroscopy enhanced by integrated plasmonic nanodimers, Opt. Express. 24 (2016) 180–190. doi:10.1364/OE.24.00A180.

[107] C. Catala, B. Mir-simon, X. Feng, C. Cardozo, N. Pazos-perez, E. Pazos, S. Gómez-de Pedro, L. Guerrini, A. Soriano, J. Vila, F. Marco, E. Garcia-rico, R.A. Alvarez-puebla, Online SERS Quantification of Staphylococcus aureus and the Application to Diagnostics in Human Fluids, Adv. Mater. Technol. 1 (2016). doi:10.1002/admt.201600163.

[108] M. Naji, A. Khetani, N. Lagali, R. Munger, H. Anis, A novel method of using hollow-

core photonic crystal fiber as a Raman biosensor, Proc. SPIE. 6865 (2008) 68650E-1-68650E–8. doi:10.1117/12.763158.

[109] X. Yang, C. Gu, F. Qian, Y. Li, J.Z. Zhang, Highly Sensitive Detection of Proteins and Bacteria in Aqueous Solution Using Surface-Enhanced Raman Scattering and Optical Fibers, Anal. Chem. 83 (2011) 5888–5894. doi:10.1021/ac200707t.

[110] N. Leopold, B. Lendl, A new method for fast preparation of highly surface-enhanced raman scattering (SERS) active silver colloids at room temperature by reduction of silver nitrate with hydroxylamine hydrochloride, J. Phys. Chem. B. 107 (2003) 5723–5727. doi:10.1021/jp027460u.

[111] A. Momenpour, P.D.A. Lima, Y.-A. Chen, C.-R. Tzeng, B.K. Tsang, H. Anis, Surface-enhanced Raman scattering for the detection of polycystic ovary syndrome, Biomed. Opt. Express. 9 (2018) 801. doi:10.1364/boe.9.000801.

[112] R. Weissleder, A clearer vision for in vivo imaging, Nat. Biotechnol. 19 (2001) 316–317. doi:10.1038/86684.

[113] A.S. Joshi, P. Singh, I. Mijakovic, Interactions of gold and silver nanoparticles with bacterial biofilms: Molecular interactions behind inhibition and resistance, Int. J. Mol. Sci. 21 (2020) 1–24. doi:10.3390/ijms21207658.

[114] E.I. Alarcon, K. Udekwu, M. Skog, N.L. Pacioni, K.G. Stamplecoskie, M. González-Béjar, N. Polisetti, A. Wickham, A. Richter-Dahlfors, M. Griffith, J.C. Scaiano, The biocompatibility and antibacterial properties of collagen-stabilized , photochemically prepared silver nanoparticles, Biomaterials. 33 (2012) 4947–4956. doi:10.1016/j.biomaterials.2012.03.033.

[115] D.R. Osmon, E.F. Berbari, A.R. Berendt, D. Lew, W. Zimmerli, J.M. Steckelberg, N. Rao, A. Hanssen, W.R. Wilson, Diagnosis and management of prosthetic joint infection: Clinical practice guidelines by the infectious diseases Society of America, Clin. Infect. Dis. 56 (2013) 1–25. doi:10.1093/cid/cis803.

[116] O. Opota, A. Croxatto, G. Prod'hom, G. Greub, Blood culture-based diagnosis of

bacteraemia: State of the art, Clin. Microbiol. Infect. 21 (2015) 313–322. doi:10.1016/j.cmi.2015.01.003.

[117] K. Kelley, A. Cosman, P. Belgrader, B. Chapman, D.C. Sullivan, Detection of Methicillin-Resistant Staphylococcus aureus by a Duplex Droplet Digital PCR Assay, J. Clin. Microbiol. 51 (2013) 2033–2039. doi:10.1128/JCM.00196-13.

[118] P.P. Banada, S. Chakravorty, D. Shah, M. Burday, F.M. Mazzella, D. Alland, Highly Sensitive Detection of Staphylococcus aureus Directly from Patient Blood, PLoS One. 7 (2012). doi:10.1371/journal.pone.0031126.

[119] E.A. Balazs, D. Watson, I.F. Duff, S. Roseman, Hyaluronic acid in synovial fluid. I. Molecular parameters of hyaluronic acid in normal and arthritic human fluids, Arthritis Rheum. 10 (1967) 357–376. doi:10.1002/art.1780100407.

[120] L.E. Guenther, B.W. Pyle, T.R. Turgeon, E.R. Bohm, U.P.W. Yss, T.A. Schmidt, J.M. Brandt, Biochemical analyses of human osteoarthritic and periprosthetic synovial fluid, Proc. Inst. Mech. Eng. Part H J. Eng. Med. 228 (2014) 127–139. doi:10.1177/0954411913517880.

[121] J. Fishovitz, J.A. Hermoso, M. Chang, S. Mobashery, Penicillin-binding protein 2a of methicillin-resistant Staphylococcus aureus, IUBMB Life. 66 (2014) 572–577. doi:10.1002/iub.1289.

[122] A. Csaki, F. Jahn, I. Latka, T. Henkel, D. Malsch, T. Schneider, K. Schröder, K. Schuster, A. Schwuchow, R. Spittel, D. Zopf, W. Fritzsche, Nanoparticle layer deposition for plasmonic tuning of microstructured optical fibers, Small. 6 (2010) 2584–2589. doi:10.1002/smll.201001071.

[123] A.C. Peacock, A. Amezcua-Correa, J. Yang, P.J.A. Sazio, S.M. Howdle, Highly efficient surface enhanced Raman scattering using microstructured optical fibers with enhanced plasmonic interactions, Appl. Phys. Lett. 92 (2008) 1–4. doi:10.1063/1.2907506.

[124] A. Amezcua-Correa, J. Yang, C.E. Finlayson, A.C. Peacock, J.R. Hayes, P.J.A.

Sazio, J.J. Baumberg, S.M. Howdle, Surface-enhanced Raman scattering using microstructured optical fiber substrates, Adv. Funct. Mater. 17 (2007) 2024–2030. doi:10.1002/adfm.200601125.

[125] Y. Guo, M.K. Khaing Oo, K. Reddy, X. Fan, Ultrasensitive Optofluidic Surface-Enhanced Raman Scattering Detection with Flow-through Multihole Capillaries, (2012) 381–388.

[126] Y. Han, S. Tan, M.K. Khaing Oo, D. Pristinski, S. Sukhishvili, H. Du, Towards full-length accumulative surface-enhanced raman Scattering-Active photonic crystal fibers, Adv. Mater. 22 (2010) 2647–2651. doi:10.1002/adma.200904192.

[127] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA. Cancer J. Clin. 68 (2018) 394–424. doi:10.3322/caac.21492.

[128] D. Jelovac, D.K. Armstrong, Recent progress in the diagnosis and treatment of Ovarian Cancer, CA. Cancer J. Clin. 61 (2011) 183–203. doi:10.3322/caac.20113.Available.

[129] R. Pokhriyal, R. Hariprasad, L. Kumar, G. Hariprasad, Chemotherapy Resistance in Advanced Ovarian Cancer Patients, Biomark. Cancer. 11 (2019) 1–19. doi:10.1177/1179299x19860815.

[130] M.X. Liu, M.K.Y. Siu, S.S. Liu, J.W.P. Yam, H.Y.S. Ngan, D.W. Chan, Epigenetic silencing of microRNA-199b-5p is associated with acquired chemoresistance via activation of JAG1-Notch1 signaling in ovarian cancer, Oncotarget. 5 (2014) 944–958. doi:10.18632/oncotarget.1458.

[131] V. Loizzi, J.K. Chan, K. Osann, F. Cappuccini, P.J. DiSaia, M.L. Berman, Survival outcomes in patients with recurrent ovarian cancer who were treated with chemoresistance assay-guided chemotherapy, Am. J. Obstet. Gynecol. 189 (2003) 1301–1307. doi:10.1067/S0002-9378(03)00629-X.

[132] T. Singh, A.S. Neal, N.A. Moatamed, S. Memarzadeh, Exploring the potential of drug response assays for precision medicine in ovarian cancer, Int. J. Mol. Sci. 22 (2021) 1–17. doi:10.3390/ijms22010305.

[133] G. Konecny, C. Crohns, M. Pegram, M. Felber, S. Lude, C. Kurbacher, I.A. Cree, H. Hepp, M. Untch, Correlation of drug response with the ATP tumorchemosensitivity assay in primary FIGO stage III ovarian cancer, Gynecol. Oncol. 77 (2000) 258–263. doi:10.1006/gyno.2000.5728.

[134] T.H. Lippert, H.J. Ruoff, M. Volm, Current status of methods to assess cancer drug resistance, Int. J. Med. Sci. 8 (2011) 245–253. doi:10.7150/ijms.8.245.

[135] G.L. Bellot, W.H. Tan, L.L. Tay, D. Koh, X. Wang, Reliability of tumor primary cultures as a model for drug response prediction: Expression profiles comparison of tissues versus primary cultures from colorectal cancer patients, J. Cancer Res. Clin. Oncol. 138 (2012) 463–482. doi:10.1007/s00432-011-1115-9.

[136] F. Caobelli, P. Alongi, L. Evangelista, M. Picchio, G. Saladini, M. Rensi, O. Geatti, A. Castello, I. Laghai, C.E. Popescu, C. Dolci, C. Crivellaro, S. Seghezzi, M. Kirienko, V. De Biasi, F. Cocciolillo, N. Quartuccio, Young AIMN Working Group, Predictive value of 18F-FDG PET/CT in restaging patients affected by ovarian carcinoma: a multicentre study, Eur. J. Nucl. Med. Mol. Imaging. 43 (2016) 404–413. doi:10.1007/s00259-015-3184-5.

[137] J.K. Schwarz, P.W. Grigsby, F. Dehdashti, D. Delbeke, The role of 18F-FDG PET in assessing therapy response in cancer of the cervix and ovaries, J. Nucl. Med. 50 (2009) 64–73. doi:10.2967/jnumed.108.057257.

[138] R. Liu, G. Zhang, Z. Yang, Towards rapid prediction of drug-resistant cancer cell phenotypes: single cell mass spectrometry combined with machine learning, Chem. Commun. 55 (2019) 616–619. doi:10.1039/c8cc08296k.

[139] M.K. Hammoud, H.K. Yosef, T. Lechtonen, K. Aljakouch, M. Schuler, W. Alsaidi, I. Daho, A. Maghnouj, S. Hahn, S.F. El-Mashtoly, K. Gerwert, Raman micro-

spectroscopy monitors acquired resistance to targeted cancer therapy at the cellular level, Sci. Rep. 8 (2018) 1–11. doi:10.1038/s41598-018-33682-7.

[140] H. Nawaz, F. Bonnier, A.D. Meade, F.M. Lyng, H.J. Byrne, Comparison of subcellular responses for the evaluation and prediction of the chemotherapeutic response to cisplatin in lung adenocarcinoma using Raman spectroscopy, Analyst. 136 (2011) 2450–2463. doi:10.1039/c1an15104e.

[141] H. Huang, H. Shi, S. Feng, W. Chen, Y. Yu, D. Lin, R. Chen, Confocal Raman spectroscopic analysis of the cytotoxic response to cisplatin in nasopharyngeal carcinoma cells, Anal. Methods. 5 (2013) 260–266. doi:10.1039/c2ay25684c.

[142] H. Moradi, A. Ahmad, D. Shepherdson, N.H. Vuong, G. Niedbala, L. Eapen, B. Vanderhyden, B. Nyiri, S. Murugkar, Raman micro-spectroscopy applied to treatment resistant and sensitive human ovarian cancer cells, J. Biophotonics. 10 (2017) 1327–1334. doi:10.1002/jbio.201600211.

[143] K.D. Steffensen, M. Waldstrøm, I. Brandslund, M. Petzold, A. Jakobsen, The prognostic and predictive value of combined HE4 and CA-125 in ovarian cancer patients, Int. J. Gynecol. Cancer. 22 (2012) 1474–1482. doi:10.1097/IGC.0b013e3182681cfd.

[144] K. Partheen, B. Kristjansdottir, K. Sundfeldt, Evaluation of ovarian cancer biomarkers HE4 and CA-125 in women presenting with a suspicious cystic ovarian mass, J. Gynecol. Oncol. 22 (2011) 244–252. doi:10.3802/jgo.2011.22.4.244.

[145] C. Théry, L. Zitvogel, S. Amigorena, Exosomes: Composition, biogenesis and function, Nat. Rev. Immunol. 2 (2002) 569–579. doi:10.1038/nri855.

[146] Z.Y.A. Elmageed, Y. Yang, R. Thomas, M. Ranjan, D. Mondal, K. Moroz, Z. Fang, B.M. Rezk, K. Moparty, S.C. Sikka, O. Sartor, A.B. Abdel-Mageed, Neoplastic reprogramming of patient-derived adipose stem cells by prostate cancer cell-associated exosomes, Stem Cells. 32 (2014) 983–997. doi:10.1002/stem.1619.

[147] M. Asare-Werehene, K. Nakka, A. Reunov, C.T. Chiu, W.T. Lee, M.R. Abedini,

P.W. Wang, D. Bin Shieh, F.J. Dilworth, E. Carmona, T. Le, A.M. Mes-Masson, D. Burger, B.K. Tsang, The exosome-mediated autocrine and paracrine actions of plasma gelsolin in ovarian cancer chemoresistance, Oncogene. 39 (2020) 1600–1616. doi:10.1038/s41388-019-1087-9.

[148] J. Crow, S. Atay, S. Banskota, B. Artale, S. Schmitt, A.K. Godwin, Exosomes as mediators of platinum resistance in ovarian cancer, Oncotarget. 8 (2017) 11917–11936. doi:10.18632/oncotarget.14440.

[149] W.X. Chen, X.M. Liu, M.M. Lv, L. Chen, J.H. Zhao, S.L. Zhong, M.H. Ji, Q. Hu, Z. Luo, J.Z. Wu, J.H. Tang, Exosomes from drug-resistant breast cancer cells transmit chemoresistance by a horizontal transfer of MicroRNAs, PLoS One. 9 (2014). doi:10.1371/journal.pone.0095240.

[150] R. Safaei, B.J. Larson, T.C. Cheng, M.A. Gibson, S. Otani, W. Naerdemann, S.B. Howell, Abnormal lysosomal trafficking and enhanced exosomal export of cisplatin in drug-resistant human ovarian carcinoma cells, Mol. Cancer Ther. 4 (2005) 1595–1604. doi:10.1158/1535-7163.MCT-05-0102.

[151] X.J. Liang, D.W. Shen, K.G. Chen, S.M. Wincovitch, S.H. Garfield, M.M. Gottesman, Trafficking and localization of platinum complexes in cisplatin-resistant cell lines monitored by fluorescence-labeled platinum, J. Cell. Physiol. 202 (2005) 635–641. doi:10.1002/jcp.20253.

[152] T. Falta, P. Heffeter, A. Mohamed, W. Berger, S. Hann, G. Koellensperger, Quantitative determination of intact free cisplatin in cell models by LC-ICP-MS, J. Anal. At. Spectrom. 26 (2011) 109–115. doi:10.1039/c0ja00047g.

[153] M. Basotra, S.K. Singh, M. Gulati, Development and Validation of a Simple and Sensitive Spectrometric Method for Estimation of Cisplatin Hydrochloride in Tablet Dosage Forms: Application to Dissolution Studies, ISRN Anal. Chem. (2013) 1–8. doi:10.1155/2013/936254.

[154] T. Jantarat, S. Chuaychob, C. Thammakhet-Buranachai, P. Thavarungkul, P.

Kanatharana, W. Srisintorn, C. Buranachai, A Label-free DNA-based Fluorescent Sensor for Cisplatin Detection, Sensors Actuators, B Chem. (2021) 128764. doi:10.1016/j.snb.2020.128764.

[155] E.H. Koritzinsky, J.M. Street, R.A. Star, P.S.T. Yuen, Quantification of Exosomes, J. Cell. Physiol. 232 (2017) 1587–1590. doi:10.1002/jcp.25387.Quantification.

[156] M. Asare-Werehene, L. Communal, E. Carmona, Y. Han, Y.S. Song, D. Burger, A.M. Mes-Masson, B.K. Tsang, Plasma gelsolin inhibits CD8þ t-cell function and regulates glutathione production to confer chemoresistance in ovarian cancer, Cancer Res. 80 (2020) 3959–3971. doi:10.1158/0008-5472.CAN-20-0788.

[157] Z.M. Zhang, S. Chen, Y.Z. Liang, Baseline correction using adaptive iteratively reweighted penalized least squares, Analyst. 135 (2010) 1138–1146. doi:10.1039/b922045c.

[158] M.H. Baik, R.A. Friesner, S.J. Lippard, Theoretical Study of Cisplatin Binding to Purine Bases: Why Does Cisplatin Prefer Guanine over Adenine?, J. Am. Chem. Soc. 125 (2003) 14082–14092. doi:10.1021/ja036960d.

[159] C. Jing, Y. Fang, Experimental (SERS) and theoretical (DFT) studies on the adsorption behaviors of l-cysteine on gold/silver nanoparticles, Chem. Phys. 332 (2007) 27–32. doi:10.1016/j.chemphys.2006.11.019.

[160] C. Gardiner, M. Shaw, P. Hole, J. Smith, D. Tannetta, C.W. Redman, I.L. Sargent, Measurement of refractive index by nanoparticle tracking analysis reveals heterogeneity in extracellular vesicles, J. Extracell. Vesicles. 3 (2014) 1–6. doi:10.3402/jev.v3.25361.

[161] J.S. Gebauer, M. Malissek, S. Simon, S.K. Knauer, M. Maskos, R.H. Stauber, W. Peukert, L. Treuel, Impact of the nanoparticle-protein corona on colloidal stability and protein structure, Langmuir. 28 (2012) 9673–9679. doi:10.1021/la301104a.

[162] M. Chisanga, H. Muhamadali, D.I. Ellis, R. Goodacre, Surface-Enhanced Raman Scattering (SERS) in Microbiology: Illumination and Enhancement of the Microbial

World, Appl. Spectrosc. 72 (2018) 987–1000. doi:10.1177/0003702818764672.

[163] G.P. Szekeres, J. Kneipp, SERS probing of proteins in gold nanoparticle agglomerates, Front. Chem. 7 (2019) 1–10. doi:10.3389/fchem.2019.00030.

[164] Y. Shen, L. Liang, S. Zhang, D. Huang, R. Deng, J. Zhang, H. Qu, S. Xu, C. Liang, W. Xu, Organelle-Targeting Gold Nanorods for Macromolecular Profiling of Subcellular Organelles and Enhanced Cancer Cell Killing, ACS Appl. Mater. Interfaces. 10 (2018) 7910–7918. doi:10.1021/acsami.8b01320.

[165] B. Hao, K. Wang, Y. Zhou, C. Sui, L. Wang, R. Bai, Z. Yang, Label-Free Detecting of the Compaction and Decompaction of ctDNA Molecules Induced by Surfactants with SERS Based on a nanoPAA-ZnCl2-AuLs Solid Substrate, ACS Omega. 5 (2020) 1109–1119. doi:10.1021/acsomega.9b03294.

[166] M.R. Abedini, P.W. Wang, Y.-F. Huang, M. Cao, C.-Y. Chou, D.-B. Shieh, B.K. Tsang, Cell fate regulation by gelsolin in human gynecologic cancers, Proc. Natl. Acad. Sci. 111 (2014) 14442–14447. doi:10.1073/pnas.

[167] M.A. Unger, H.-P. Chou, T. Thorsen, A. Scherer, S.R. Quake, Monolithic Microfabricated Valves and Pumps by Multilayer Soft Lithography, (n.d.).

[168] H. Gong, A.T. Woolley, G.P. Nordin, High density 3D printed microfluidic valves, pumps, and multiplexers, Lab Chip. 16 (2016) 2450–2458. doi:10.1039/c6lc00565a.

[169] W.H. Grover, M.G. Von Muhlen, S.R. Manalis, Teflon films for chemically-inert microfluidic valves and pumps, Lab Chip. 8 (2008) 913–918. doi:10.1039/b800600h.

[170] J. Kim, M. Kang, E.C. Jensen, R.A. Mathies, Lifting gate PDMS microvalves and pumps for microfluidic control, Anal. Chem. 84 (2013) 2067–2071. doi:10.1021/ac202934x.Lifting.

[171] S. Haefner, R. Koerbitz, P. Frank, M. Elstner, A. Richter, High Integration of

Microfluidic Circuits Based on Hydrogel Valves for MEMS Control, Adv. Mater. Technol. 3 (2018) 1–10. doi:10.1002/admt.201700108.

[172] N.S. Satarkar, W. Zhang, R.E. Eitel, J.Z. Hilt, Magnetic hydrogel nanocomposites as remote controlled microfluidic valves, Lab Chip. 9 (2009) 1773–1779. doi:10.1039/b822694f.

[173] K. Pitchaimani, B.C. Sapp, A. Winter, A. Gispanski, T. Nishida, Z. Hugh Fan, Manufacturable plastic microfluidic valves using thermal actuation, Lab Chip. 9 (2009) 3082–3087. doi:10.1039/b909742b.

[174] W.Y. Zhang, J.P. Labukas, S. Tatic-Lucic, L. Larson, T. Bannuru, R.P. Vinci, G.S. Ferguson, Novel room-temperature first-level packaging process for microscale devices, Sensors Actuators, A Phys. 123–124 (2005) 646–654. doi:10.1016/j.sna.2005.03.008.

[175] G. Guan, L. Wu, A.A. Bhagat, Z. Li, P.C.Y. Chen, S. Chao, C.J. Ong, J. Han, Spiral microchannel with rectangular and trapezoidal cross-sections for size based particle separation, Sci. Rep. 3 (2013) 1–9. doi:10.1038/srep01475.

[176] B.L. Khoo, M.E. Warkiani, D.S.W. Tan, A.A.S. Bhagat, D. Irwin, D.P. Lau, A.S.T. Lim, K.H. Lim, S.S. Krisna, W.T. Lim, Y.S. Yap, S.C. Lee, R.A. Soo, J. Han, C.T. Lim, Clinical validation of an ultra high-throughput spiral microfluidics for the detection and enrichment of viable circulating tumor cells, PLoS One. 9 (2014) 1–7. doi:10.1371/journal.pone.0099409.

[177] M.E. Warkiani, B.L. Khoo, D.S.W. Tan, A.A.S. Bhagat, W.T. Lim, Y.S. Yap, S.C. Lee, R.A. Soo, J. Han, C.T. Lim, An ultra-high-throughput spiral microfluidic biochip for the enrichment of circulating tumor cells, Analyst. 139 (2014) 3245–3255. doi:10.1039/c4an00355a.

[178] L. Wu, G. Guan, H.W. Hou, A.A.S. Bhagat, J. Han, Separation of leukocytes from blood using spiral channel with trapezoid cross-section, Anal. Chem. 84 (2012) 9324–9331. doi:10.1021/ac302085y.

[179] W. Lee, D. Kwon, W. Choi, G.Y. Jung, A.K. Au, A. Folch, S. Jeon, 3D-Printed micro fluidic device for the detection of pathogenic bacteria using size-based separation in helical channel with trapezoid cross-section, Sci. Rep. 5 (2015) 1–7. doi:10.1038/srep07717.

[180] J.H. Lee, S.K. Lee, J.H. Kim, J.H. Park, Separation of particles with bacterial size range using the control of sheath flow ratio in spiral microfluidic channel, Sensors Actuators, A Phys. 286 (2019) 211–219. doi:10.1016/j.sna.2018.12.047.

[181] M.R. Condina, B.A. Dilmetz, S. Razavi Bazaz, J. Meneses, M. Ebrahimi Warkiani, P. Hoffmann, Rapid separation and identification of beer spoilage bacteria by inertial microfluidics and MALDI-TOF mass spectrometry, Lab Chip. 19 (2019) 1961–1970. doi:10.1039/c9lc00152b.

[182] N. Nivedita, P. Ligrani, I. Papautsky, Dean Flow Dynamics in Low-Aspect Ratio Spiral Microchannels, Sci. Rep. 7 (2017) 1–10. doi:10.1038/srep44072.

[183] H. Lim, J. Nam, S. Shin, Lateral migration of particles suspended in viscoelastic fluids in a microchannel flow, Microfluid. Nanofluidics. 17 (2014) 683–692. doi:10.1007/s10404-014-1353-7.

[184] N. Xiang, Z. Ni, H. Yi, Concentration-controlled particle focusing in spiral elasto-inertial microfluidic devices, Electrophoresis. 39 (2018) 417–424. doi:10.1002/elps.201700150.

[185] T. Trantidou, Y. Elani, E. Parsons, O. Ces, Hydrophilic surface modification of pdms for droplet microfluidics using a simple, quick, and robust method via pva deposition, Microsystems Nanoeng. 3 (2017). doi:10.1038/micronano.2016.91.

[186] S. Pavan, F. Berti, Short peptides as biosensor transducers, Anal. Bioanal. Chem. 402 (2012) 3055–3070. doi:10.1007/s00216-011-5589-8.

[187] H. Poblete, A. Agarwal, S.S. Thomas, C. Bohne, R. Ravichandran, J. Phospase, J. Comer, E.I. Alarcon, New Insights into Peptide-Silver Nanoparticle Interaction: Deciphering the Role of Cysteine and Lysine in the Peptide Sequence, Langmuir.

32 (2016) 265–273. doi:10.1021/acs.langmuir.5b03601.

[188] Z. Khatoon, I. Guzmán-Soto, C.D. Mctiernan, C. Lazurko, F. Simpson, L. Zhang, D. Cortes, T.F. Mah, M. Griffith, E.I. Alarcon, Nanoengineering the surface of corneal implants: Towards functional anti-microbial and biofilm materials, RSC Adv. 10 (2020) 23675–23681. doi:10.1039/d0ra03659e.

[189] M.K. Hossain, H.Y. Cho, K.J. Kim, J.W. Choi, In situ monitoring of doxorubicin release from biohybrid nanoparticles modified with antibody and cell-penetrating peptides in breast cancer cells using surface-enhanced Raman spectroscopy, Biosens. Bioelectron. 71 (2015) 300–305. doi:10.1016/j.bios.2015.04.053.

[190] C. Shi, X. Cao, X. Chen, Z. Sun, Z. Xiang, H. Zhao, W. Qian, X. Han, Intracellular surface-enhanced Raman scattering probes based on TAT peptide-conjugated Au nanostars for distinguishing the differentiation of lung resident mesenchymal stem cells, Biomaterials. 58 (2015) 10–25. doi:10.1016/j.biomaterials.2015.04.010.

[191] F. Tanvir, A. Yaqub, S. Tanvir, W.A. Anderson, Poly-l-arginine coated silver nanoprisms and their anti-bacterial properties, Nanomaterials. 7 (2017). doi:10.3390/nano7100296.

[192] K. Yuan, Q. Mei, X. Guo, Y. Xu, D. Yang, B.J. Sánchez, B. Sheng, C. Liu, Z. Hu, G. Yu, H. Ma, H. Gao, C. Haisch, R. Niessner, Z. Jiang, H. Zhou, Antimicrobial peptide based magnetic recognition elements and Au@Ag-GO SERS tags with stable internal standards: A three in one biosensor for isolation, discrimination and killing of multiple bacteria in whole blood, Chem. Sci. 9 (2018) 8781–8795. doi:10.1039/c8sc04637a.

[193] D. Selvamuthu, D. Das, Introduction to statistical methods, design of experiments and statistical quality control, 2018. doi:10.1007/978-981-13-1736-1.

[194] Y. Hu, X. Lu, Rapid Detection of Melamine in Tap Water and Milk Using Conjugated “One-Step” Molecularly Imprinted Polymers-Surface Enhanced Raman Spectroscopic Sensor, J. Food Sci. 81 (2016) N1272–N1280. doi:10.1111/1750-

3841.13283.

[195] P. Zahedi, M. Ziaee, M. Abdouss, A. Farazin, B. Mizaikoff, Biomacromolecule template-based molecularly imprinted polymers with an emphasis on their synthesis strategies: a review, Polym. Adv. Technol. 27 (2016) 1124–1142. doi:10.1002/pat.3754.

[196] M. Dinc, C. Esen, B. Mizaikoff, Recent advances on core–shell magnetic molecularly imprinted polymers for biomacromolecules, TrAC - Trends Anal. Chem. 114 (2019) 202–217. doi:10.1016/j.trac.2019.03.008.

[197] National Institutes of Health, Evidence-based Methodology Workshop on Polycystic Ovary Syndrome December 3-5, 2012.

[198] Y. Urata, R. Salehi, P.D.A. Lima, Y. Osuga, B.K. Tsang, Neuropeptide y regulates proliferation and apoptosis in granulosa cells in a follicular stage-dependent manner, J. Ovarian Res. 13 (2020) 1–11. doi:10.1186/s13048-019-0608-z.

[199] Y. Li, Y. Chen, L. Huang, B. Lou, G. Chen, Creating BHb-imprinted magnetic nanoparticles with multiple binding sites, Analyst. 142 (2017) 302–309. doi:10.1039/c6an02121b.

[200] S. Solorio-Fernández, J.A. Carrasco-Ochoa, J.F. Martínez-Trinidad, A review of unsupervised feature selection methods, Artif. Intell. Rev. 53 (2020) 907–948. doi:10.1007/s10462-019-09682-y.

[201] H. Sanz, C. Valim, E. Vegas, J.M. Oller, F. Reverter, SVM-RFE: Selection and visualization of the most relevant features through non-linear kernels, BMC Bioinformatics. 19 (2018) 1–18. doi:10.1186/s12859-018-2451-4.

[202] B. Kampe, S. Kloß, T. Bocklitz, P. Rösch, J. Popp, Recursive feature elimination in Raman spectra with support vector machines, Front. Optoelectron. 10 (2017) 273–279. doi:10.1007/s12200-017-0726-4.

[203] J.N. Taylor, K. Mochizuki, K. Hashimoto, Y. Kumamoto, Y. Harada, K. Fujita, T.

Komatsuzaki, High-Resolution Raman Microscopic Detection of Follicular Thyroid Cancer Cells with Unsupervised Machine Learning, J. Phys. Chem. B. 123 (2019) 4358–4372. doi:10.1021/acs.jpcb.9b01159.

[204] J. Chen, Y. Zeng, Y. Li, G. Bin Huang, Unsupervised feature selection based extreme learning machine for clustering, Neurocomputing. 386 (2020) 198–207. doi:10.1016/j.neucom.2019.12.065.

[205] D. Song, Y. Chen, J. Li, H. Wang, T. Ning, S. Wang, A graphical user interface (NWUSA) for Raman spectral processing, analysis and feature recognition, J. Biophotonics. 14 (2021) 1–12. doi:10.1002/jbio.202000456.

[206] B.R. Masters, Superresolution optical microscopy: the quest for enhanced resolution and contrast, 2020. http://link.springer.com/10.1007/978-3-030-21691-7.

[207] E.F. Fornasiero, S.O. Rizzoli, Super-Resolution Microscopy Techniques in the Neurosciences, 2015. doi:10.1117/1.jbo.20.1.019901.

[208] M.G.L. Gustafsson, L. Shao, P.M. Carlton, C.J.R. Wang, I.N. Golubovskaya, W.Z. Cande, D.A. Agard, J.W. Sedat, Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination, Biophys. J. 94 (2008) 4957–4970. doi:10.1529/biophysj.107.120345.

[209] J.-H. Han, N.-W. Yoo, J.-H. Kang, B.-K. Ju, M.-C. Park, Optimization of structured illumination microscopy with designing and rotating a grid pattern using a spatial light modulator, Opt. Eng. 58 (2019) 1. doi:10.1117/1.oe.58.9.094102.

[210] Z. Fu, Q. Geng, J. Chen, L.-A. Chu, A.-S. Chiang, S.-C. Chen, Light field microscopy based on structured light illumination, Opt. Lett. 46 (2021) 3424. doi:10.1364/ol.428562.

[211] M. Lachetta, H. Sandmeyer, A. Sandmeyer, J.S.A. Esch, T. Huser, M. MÃller, Simulating digital micromirror devices for patterning coherent excitation light in structured illumination microscopy, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 379 (2021). doi:10.1098/rsta.2020.0147.

[212] H. Chen, S. Wang, Y. Zhang, Y. Yang, H. Fang, S. Zhu, X. Yuan, Structured illumination for wide-field Raman imaging of cell membranes, Opt. Commun. 402 (2017) 221–225. doi:10.1016/j.optcom.2017.06.021.

[213] M. Wang, C. Zhang, S. Yan, T. Chen, H. Fang, X. Yuan, Wide-Field Super-Resolved Raman Imaging of Carbon Materials, ACS Photonics. 8 (2021) 1801–1809. doi:10.1021/acsphotonics.1c00392.

[214] S. Schlücker, SERS Microscopy: Nanoparticle Probes and Biomedical Applications, Surf. Enhanc. Raman Spectrosc. Anal. Biophys. Life Sci. Appl. (2010) 263–283. doi:10.1002/9783527632756.ch12.