ICFO – Institut de Ciències Fotòniques

UPC – Universitat Politècnica de Catalunya

Detection of particles, bacteria and viruses using consumer optoelectronic components

Rubaiya Binte Hussain

Thesis advisor: Prof. Valerio Pruneri

PhD Thesis- 2022

To my loving sister Eva… the one I lost to patriarchy…

ABSTRACT

Over the past few decades, various novel optical technologies for particle detection and characterization have been developed, based on the unprecedented interaction of light with matter. Light scattering, a fundamental phenomenon responsible for many physical effects, such as the blue colour of the sky, can be exploited to create essential devices. One such example of this is particle size analysers (PSA). These devices have gained popularity in industry for counting particles and determining their size distribution, which is crucial for monitoring and controlling many production processes and for biological and environmental applications. Particularly relevant for this thesis is the ability to count bacteria in drinking and recreational water or viral particles that can cause infectious diseases. One of the most common methods for such applications is flow-cytometry, which uses fluorescence, as well as scattering, to count or analyse particles. This approach is highly sensitive and specific due to the use of fluorescent markers to tag the target of interest. However, the traditional methods used for particle size and count analysis rely on large and expensive devices, often requiring complicated sample preparation, a dedicated laboratory and skilled staff.

The focus of this thesis is on the design, development and validation of two novel photonic sensors for the detection and characterisation of industrial and biological samples. The first one is a PSA in a collimated beam configuration using an innovative angular spatial filter, and a consumer electronic camera similar to that used in a smartphone. The small form factor angular spatial filter allows for the collection of diffused light from particles up to predefined discrete angles. By using angularly resolved scattering images acquired by the camera, a machine learning (ML) algorithm predicts the volume median diameter of the particles. Our system has achieved a mean absolute percentage error of only 0.72% for spherical particles in solution with sizes greater than 10 µm and at concentrations up to 40 mg mL^{-1}. Compared to traditional laser diffraction systems, the proposed PSA is an order of magnitude smaller in size, weight and cost, and offers a promising approach to online industrial process monitoring.

As light scattering is influenced by factors other than particle size, including shape, refractive index contrast and suspension concentration, the PSA can also be employed in biological applications. To this end, the second part of the thesis aims to optimise the PSA for the measurement of small (< 10 µm) particles such as microorganisms. The results demonstrate that the modified PSA in combination with ML is able to accurately classify different types of bacteria (*Escherichia coli* and *Enterococcus sp.*) and distinguish them from silica beads of comparable sizes, with an accuracy of 89%. Moreover, it can detect the concentration of bacteria in water with a limit of detection (LOD) of approximately 10^5 cells mL^{-1}.

The final part of the thesis is dedicated to the development of a low-cost, portable optical biosensor for the specific detection of particles smaller than bacteria, such as viruses (< 1 µm). The proposed system, which we have called flow virometry reader (FVR), is a modification of a flow cytometer and relies on measuring light emissions from fluorescent antibodies that bind to specific viral particles. An LOD of 3,834 copies mL^{-1} for SARS-CoV-2 in saliva can be achieved with the device. The FVR is clinically validated using 54 saliva samples in a blind test, with high sensitivity and specificity of 91.2% and 90%, respectively. These findings suggest that the FVR has the potential to be a highly viable alternative to current diagnostic methods for pandemic events, as it is faster (< 30 min) and less expensive than PCR tests, while being more sensitive than today's COVID-19 rapid antigen tests.

The photonic sensing technologies developed in the thesis show significant potential for use in a wide range of applications, including:

- particulate air pollution, causing cardiovascular and respiratory problems
- particulate water pollution, which affects the ecosystems of rivers, lakes and oceans
- total bacterial count in environmental or bathing water
- viral pandemics

The technologies are particularly appealing in countries with limited resources due to their simple design, portability, short-time-to-result and affordability, as well as the fact that they do not require a specialised laboratory or trained personnel to operate them.

RESUMEN

En las últimas décadas, se han desarrollado nuevas tecnologías ópticas para la detección y caracterización de partículas, basadas en la interacción luz-materia. La dispersión de la luz, un fenómeno fundamental responsable de varios efectos físicos, como el color azul del cielo, puede ser utilizado para crear dispositivos esenciales. Un ejemplo de estos dispositivos son los analizadores de tamaño para partículas (siglas PSA en inglés). Estos dispositivos han adquirido popularidad en la industria al poder determinar el número de partículas y la distribución de tamaños de estas, un punto crucial para monitorizar y controlar procesos de producción y para aplicaciones biológicas y medioambientales. En esta tesis, es particularmente importante la habilidad de contar bacterias en bebidas y aguas de uso recreativo o partículas víricas que puedan causar enfermedades infecciosas. Uno de los métodos más comunes para esta clase de aplicaciones es la citometría de flujo, la cual usa fluorescencia, así como la dispersión, para contar o analizar las partículas. Este método es muy sensible y específico dada la utilización de marcadores fluorescentes para etiquetar el objeto de interés. No obstante, los métodos utilizados tradicionalmente para determinar el tamaño de las partículas y hacer un recuento de estas requieren de equipos costosos y de grandes dimensiones, a veces requiriendo de preparaciones complejas en las muestras, laboratorios dedicados y personal experimentado.

El objetivo de esta tesis es el diseño, desarrollo y validación de dos nuevos sensores fotónicos para la detección y caracterización de muestras industriales y biológicas. El primero es un PSA en configuración de haz colimado que usa un innovador filtro espacial angular y una cámara electrónica similar a la usada en móviles. El pequeño factor de tamaño del filtro angular espacial permite la detección de la luz difusa de las partículas hasta ángulos discretos predefinidos. A partir del uso de imágenes difusas angularmente resueltas obtenidas por la cámara, un algoritmo de aprendizaje automático, *machine learning* (ML) en inglés, puede predecir la mediana del diámetro del volumen de las partículas. Nuestro sistema ha conseguido un error absoluto medio porcentual de solamente un 0.72% para partículas esféricas en disoluciones con tamaños superiores a 10 µm y concentraciones de hasta 40 mg mL^{-1}. En comparación a sistemas tradicionales de difracción láser, el propuesto PSA es un orden de magnitud más pequeño en tamaño, peso y coste, y ofrece un enfoque prometedor para la supervisión online de procesos industriales.

Dado que la difusión de luz depende de más factores aparte del tamaño de la partícula, incluyendo la forma, el contraste del índice de refracción y la suspensión de la concentración, el PSA también puede ser empleado en aplicaciones biológicas. Con este objetivo, la segunda parte de la tesis busca optimizar el PSA para la medida de partículas pequeñas (< 10 µm) como microorganismos. Los resultados demuestran que el PSA modificado en combinación con ML es capaz de clasificar con exactitud diferentes tipos de bacterias (*Escherichia coli* y *Enterococcus sp.*) y diferéncialas de partículas de silicio con tamaños similares, con una precisión del 89%. Además, puede detectar una concentración de bacterias en agua con un límite de detección (LOD en inglés) de aproximadamente 10^5 células mL^{-1}.

La parte final de tesis está dedicada al desarrollo de un biosensor óptico de bajo coste y portátil para la detección especifica de partículas más pequeñas que bacterias, como virus (< 1 µm). El sistema propuesto, el cual hemos llamado *flow virometry reader* (FVR), es una modificación de un citómetro de flujo y se basa en la medida de emisiones de luz provenientes de anticuerpos fluorescentes que son unidos a partículas virales específicas. Con este dispositivo se puede

conseguir un LOD de 3,834 copias mL^{-1} para el SARS-CoV-2 en saliva. El FVR ha sido validado clínicamente usando 54 muestras de saliva en un test a ciegas, con una sensibilidad y especificidad del 91.2% y 90%, respectivamente. Estos hallazgos sugieren que el FVR tiene el potencial de ser una alternativa viable a los métodos de diagnóstico actuales en escenarios de pandemias, pues es rápido (< 30 min) y menos costoso que los test por PCR, mientras que es más sensible que los actuales test de antígenos para COVID-19.

Las tecnologías de detección fotónicas desarrolladas en esta tesis muestran un potencial significativo para su uso en un amplio rango de aplicaciones, incluyendo:

- contaminación de aire por partículas, causantes de problemas cardiovasculares y respiratorios
- contaminación de agua por partículas, el cual afecta a ecosistemas como ríos, lagos y océanos
- recuento total de bacterias en aguas de baño o ambientales
- pandemias víricas.

Estas tecnologías son particularmente atractivas en países con recursos limitados, dado sus simples diseños, portabilidad, el poco tiempo de espera para obtener resultados y asequibilidad, así como el hecho de que estos no requieren un laboratorio especializado o un personal cualificado para operar con ellas.

FUNDING

ACKNOWLEDGEMENTS

The acknowledgement section of a thesis is probably the most difficult part to write as it brings back all the memories of a PhD journey. Looking back, I can certainly say that the last five years of my life at ICFO have been the best and the happiest part of my life, so far, despite the numerous struggles and disappointments that every PhD student has to go through.

The person who gave me this opportunity to be a part of an amazing group and work on an interesting project is my mentor and supervisor, Valerio Pruneri. I am immensely grateful to him for always guiding me, making it possible for me to explore new scientific territories and dedicating his time to solving any problem that I faced, no matter how small it was. Once he told me that the job of a mentor is to bring out the best in his students, and although in this process he challenged me numerous times, I believe he did manage to be successful there. I thank him for his patience in all those times that I have been difficult and for all those moments he gave me confidence by making me focus on my achievements rather than on my failures.

In this journey, I consider myself very lucky to have had more than one mentor to whom I could turn for advice on not only technical issues in the lab but also any other problems I encountered during my PhD. Our incredibly knowledgeable mamma, Vittoria, has been the greatest support during one of the most difficult times of my PhD and I would not have been able to overcome those challenges without her. I still bother her with my problems every now and then, but I bet she loves me despite all the troubles! Thank you Vic for taking care of me all these years.

During my last PhD project, I have had the opportunity to work closely with two very talented Postdocs of the group, Ewelina and Alfredo. I am grateful to them for everything they have taught me, for being extremely collaborative, and for always helping me in any way they could. Over the years, we've become good friends and, above all, a great team, sharing moments of "euphoria" and extreme frustration while pulling off some fantastic work together.

I would also like to express my gratitude to Angel and Merche from Biolab for helping me in my experiments, Maria for helping me numerous times with image processing, Xavi from the Mechanical workshop for being very prompt in making the parts I needed for my optical setup, Ingrid and Anne from HR for helping me with all the bureaucratic work, Santi and Magda from purchasing, Alina, Jordi, Tomas, Andrea and all those from the different units of ICFO who have helped me over the years and with their work make ICFO a great institute for researchers.

I feel privileged to have been part of the Optoelectronics group, which more than just a group of colleagues has become more of a big family, always trying to help and support one another through thick and thin. I have shared beautiful moments with the Opto family, engaging in discussions and learning from them about their beliefs, their culture and their country. I thank each one of you, the new and the old members (especially Pedro, Marc, Cedric, Alican, Rafael, Luc, Kavitha, and Rinu), for directly or indirectly helping me in my work and making my PhD time enjoyable.

Thanks to my officemates, the old team JR, JM, Vahagn and the new team Dani, Nestor, Javi and JM's soul (as he refuses to leave this office), for making Office 270 the most happening office in the group, for being patient with me when I went a little crazy, for keeping me entertained and for sharing these fun-filled 4.5 years of my PhD. Thanks also to the post-pandemic lunch and coffee buddies Dani, Javi and David for the very serious discussions we had and continue to have about ducks, turtles, and what not!

This group has also given me some of my very good friends who, outside ICFO, became my very own support system. I would like to thank *La Familia* for keeping me sane all these years and standing by my side, especially during those days when I was struggling the most. I could never have imagined that thousands of miles away from home, I would find friends who would become more than my family. Without JM's classic COVID parties, those zoom exercise sessions during the lockdown that I never joined, our meetings in the park and our WhatsApp calls and chats, I do not know how I would have survived the pandemic all alone.

Thanks to my Parchis team Miri, Seba and Laura, for being there for me, especially during those tough times, and always wanting wholeheartedly to see me happy. I am not so sure about Seba though. A special thanks to my Marchena for being that friend with whom I could share anything, even my deepest darkest secrets, without any fear of ever being judged.

JR, JM, Sofi, Manu, Noslen, Alberto & Kevin, thanks to all of you for supporting me always and for being an important part of my life. Nosberto, an extra thanks for giving me my little Seba, who brightens my day with that sweet innocent smile.

I have had the privilege to know some amazingly strong and inspiring women, Sarah, Silvana, Yu Bi, Nos, Miri, Jyoti from whom I have learnt many important life lessons. I thank you for teaching me to be strong and resilient in life.

Thanks to Roland, Lisa, Laurent, Gloria, Judith, Ainhoa, Costi, Mari, Candela and many others from and outside ICFO with whom I have shared a vermut or two and who have made my stay in Barcelona even more enjoyable and my PhD a little less painful.

I would like to thank Ammu and Abbu (my parents), for paving the path for me with their innumerable sacrifices to get to this stage in my life. Even though I come from a conservative and mostly male dominating society, my parents have raised me as an equal to my brother and have always made me believe that being born a girl does not make me any less capable, and there is absolutely nothing that I cannot achieve in life. Thank you for having faith in me.

Thanks also to my niece Maisha and nephew Rizwan for always telling me how proud they are of me and cheering me on in all of my achievements. I would like to thank my extended family and my in-laws, especially Ma (my mother-in-law), who always encourages me to pursue my career, and who was the one to calm me down during those stressful days of my PhD.

Last but not the least, I take this opportunity to thank the one person who has been there for me and with me in the last nine years of my life abroad, my partner in crime and life, Fahd. Even though we struggled immensely in the past five years of our long-distance relationship and then marriage, you have managed to stay by my side always. You have been my greatest support, my best friend, my travel partner, my therapist, my punch bag and much more. I thank you for being there always, for your patience during this PhD journey of mine, for consoling me when I failed, for celebrating with me when I succeeded, for sharing with me all the bittersweet moments and experiences in life, and for loving me unconditionally.

LIST OF PUBLICATIONS

Journal Articles

- Hussain, R., Alican Noyan, M., Woyessa, G. *et al.* **An ultra-compact particle size analyser using a CMOS image sensor and machine learning**. *Light Sci Appl* **9,** 21 (2020).

- Rubaiya Hussain, Alfredo E. Ongaro, Maria L. Rodriguez de la Concepción, et al, **Small form factor flow virometer for SARS-CoV-2**, *Biomed. Opt. Express* 13, 1609-1619 (2022).

Patent Application

- Alfredo Ongaro, Rubaiya Hussain, Valerio Pruneri, Lluis Torner, **ID0221-00: Integrated Optofluidic System For Aerosol Analysis.**

Journal Article in Preparation

- R. Hussain, E. Wajs, A. Ongaro, A. Noyan, V. Pruneri, **Identification of bacterial pathogens: Escherichia Coli and Enterococcus using light scattering and machine learning.**

Relevant Presentations

- R. Camphausen, R. Hussain, R. Terborg, L. Duempelmann. A. Cuevas, E. Wajs, S. Diefenbach, V. Pruneri. **Detection of particles, micro-organisms and biomarkers using CMOS image sensors**. NanoBio&Med2019 International Conference, Barcelona, Spain, 2019. (Invited Talk)

- R. Hussain, R. Camphausen, R. Terborg, L. Duempelmann. A. Cuevas, V. Pruneri. **Large field of view imaging with classical and quantum light**. International conference by EOS European Optical Society: Optical Microsystems OµS19 Anacapri, Island of Capri, Italy, 2019. (Oral Presentation)

- R. Hussain, P. Martinez, T. Hazlehurst, F. Mahdi, V. Pruneri. **Compact and cost-effective granulometer for particle size monitoring in industrial applications.** SPIRE Workshop, Tackling the future of plant operation, Barcelona, Spain, 2017. (Poster presentation)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

In recent years, particle detection and characterisation have become crucial in many fields of industrial research. For example, in the biopharmaceutical industry manufacturing processes are required to comply with strict regulations regarding particulates to ensure product safety, quality and efficacy[1]. Characterising particles is also critical in food processing for managing the taste, texture, appearance, stability, processability, and functionality of the final product. Other examples of industrial processes that require the ability to determine particle characteristics include catalysis, adhesion, grinding and electro-deposition, among many others[3]. Moreover, in medical diagnostics, the timely detection and identification of particles such as bacteria and viruses is critical in treating many life-threating diseases[4]. Another important application involves monitoring environmental pollutants such as microplastic[5] in sea water or airborne particles[6] that cause respiratory diseases.

With the advent of optical sensors, conventional particle characterisation methods have been gradually replaced with light-based techniques in most of the above-mentioned industries. The unique interaction of light with particles led to the development of these sensors decades ago, even before the invention of the first laser in the 60s. By using the intrinsic properties of laser light, the precision and accuracy of optical sensors improved. Since then, these sensors have experienced an exponential growth in research and development, and have gained much popularity due to their inherent benefits over traditional analytical approaches. For example, optical sensors enable real-time fast detection which is also cost-effective and highly sensitive[7]. In addition, most of the time they are non-invasive as they allow complete sample recovery after measurement. This is particularly important in industries where material costs are continuously monitored and must be kept to the minimum.

The main objective of particle characterisation is to determine information regarding a particle's physical and geometrical properties, such as counts, size and distribution, shape, specific surface area etc[3]. Among these, the characterisation of a particle's size is the most important because the behaviour of a particulate system, as well as many of its physical properties, is highly impacted by its size. The dissolution rate of active components in drugs, stability of emulsions and suspensions (e.g., sediments, paints), powder flowability, viscosity, texture and feel (e.g., food products), and finish of paints and coatings are all affected by particle size distribution (PSD). As a result, in many industrial manufacturing processes, analysing the PSD of samples is crucial in controlling the quality of the end product and helping manufacturers to comply with industry standards and regulations.

Traditional methods for determining PSD include sieving and sedimentation, which have now been substituted in industries worldwide with state-of-the-art laser diffraction (LD)[8] based systems. LD devices probe the radial distribution of diffracted or scattered light at a defined

distance from the sample. The size distribution is then estimated from the scattering pattern using appropriate optical models. Commercially available LD devices gained popularity due to their broad dynamic range, robustness, high accuracy, repeatability and reproducibility. They are, however, large (each dimension is on the order of half a meter), heavy (tens of kilograms), and costly (often costing a hundred thousand euros or more). Furthermore, their complexity, along with the fact that they frequently demand maintenance and highly trained personnel, renders them impractical in most online industrial applications, which necessitate the installation of probes in processing environments, often at numerous sites.

In addition to particle size determination, particle counting is essential in many biological applications, such as monitoring the presence of airborne microorganisms in controlled environments[9], determining contamination levels like bacterial counts in drinking water[10], and detecting and quantifying viral particles that cause infectious diseases[11], among many others. The most commonly used optical technique for particle counting and characterisation in suspension is flow-cytometry. Not only can this method detect light scattering from particles, but it can also detect fluorescence. Fluorescence is caused by the molecular absorption of incident high energy photons, which leads to the spontaneous emission of photons with lower energy, i.e., longer wavelength, than the incident photons. Hence, in order to detect fluorescence, the particles must either be fluorescent or conjugated with appropriate fluorophores. Fluorescence detection is a highly specific and sensitive method for identifying and distinguishing target particles from other particles and can be particularly useful in characterising diseases in clinical settings. Although conventional flow-cytometers are multifunctional devices capable of rapid and reliable particle analysis, they are large and expensive, with a price typically ranging from a hundred to five hundred thousand euros. Furthermore, their operation necessitates trained personnel, complex sample pre-treatment involving expensive reagents and large sample volumes. They are not suitable for point-of-care (POC) applications, and research institutions and hospitals are often required to share a single system among different departments due to their high price. Therefore, their implementation in many rural areas or developing countries with limited resources becomes extremely challenging.

1.1 Objectives of the Thesis

The discussion in the previous section demonstrates the need to develop compact, portable, easy-to-use and cost-effective systems for particle characterisation. Therefore, the main aim of this thesis is to develop novel optical sensors of small form factor using consumer electronic components for the non-invasive detection of particles, including bacteria and viruses. The specific objectives involve:

  i. Development of a miniaturised particle size analyser (PSA) based on light scattering and an ad hoc machine learning algorithm to measure particles in the size range from 10 µm to 100 µm.
 ii. Optimisation of the PSA for small particle measurements, targeting between 1 µm and 10 µm, to demonstrate its applicability in detecting microorganisms like *Escherichia coli (E. coli)* and *Enterococcus sp.* in water.
iii. Development of an application-specific flow-virometry POC reader that utilises the fluorescence measurement feature of a flow-cytometer in order to detect and quantify particles smaller than 1 µm, such as SARS-CoV-2 (Severe acute respiratory syndrome coronavirus type 2) in saliva.

1.2 Thesis Outline

The thesis is organised into five chapters:

This chapter describes the importance of particle detection and characterisation for industrial, as well as biological, applications and presents the main Thesis' objectives.

Chapter 2 presents the notion of particles, as well as various non-optical and optical particle characterisation techniques, with a focus on well-known and commercial state-of-the-art methods that are currently used in the industry. Since the optical methods presented in this chapter are dependent on light scattering, some theoretical background and modelling of light scattering are also included.

Chapter 3 introduces the concept of a light scattering based compact and low-cost PSA. This device uses consumer electronic components, a simple light emitting diode (LED) and a single metal-oxide-semiconductor (CMOS) image sensor, similar to those used in smart phones, as well as a custom machine learning model to predict the median particle size. This chapter also explains how the system is validated by measuring particles with various size distributions ranging from 10 µm to 100 µm at several concentrations.

To measure particles smaller than 1 µm, the developed PSA requires design changes, which is the focus of Chapter 4. This chapter also presents the preliminary results from measuring different concentrations of *E. coli* and *Enterococcus sp.*in water as an application case study, demonstrating how the PSA can classify *E. coli*, *Enterococcus sp.,* and other particles while also predicting bacterial concentrations in water using two different machine learning models.

For the specific and sensitive detection of particles smaller than 1 µm, like viruses, direct light scattering measurement is insufficient. Therefore, Chapter 5 introduces the notion of a small-form-factor flow-virometry reader (FVR) for POC applications, which employs a small microfluidic chip and a laser-pumped optical head to detect the presence of viruses tagged with fluorescent antibodies. The FVR's performance is assessed using clinical samples of SARS-CoV-2 in saliva, as described in the chapter. The detection limit and sensitivity obtained from the measurements are also presented.

Finally, Chapter 6 summarises the conclusions of the thesis and provides an outlook for future work

1.3 References

1. Shekunov, B. Y., Chattopadhyay, P., Tong, H. H. Y. & Chow, A. H. L. Particle size analysis in pharmaceutics: Principles, methods and applications. *Pharm. Res.* **24**, 203–227 (2007).

2. Robins, M. M. Particle Size Analysis in Food. *Encycl. Anal. Chem.* (2006) doi:10.1002/9780470027318.A1021.

3. Xu, R. Light scattering: A review of particle characterization applications. *Particuology* vol. 18 11–21 (2015).

4. Castillo-Henríquez, L. *et al.* Biosensors for the detection of bacterial and viral clinical pathogens. *Sensors (Switzerland)* **20**, 1–26 (2020).

5. Parrish, K. & Fahrenfeld, N. L. Microplastic biofilm in fresh- and wastewater as a function of microparticle type and size class. *Environ. Sci. Water Res. Technol.* **5**, 495–505 (2019).

6. Harrison, R. M. Key pollutants—airborne particles. *Sci. Total Environ.* **334**–**335**, 3–8

(2004).

7. Narayanaswamy, R. & Wolfbeis, O. S. *Optical Sensors*. *Medical Devices and Systems* vol. 1 (Springer Berlin Heidelberg, 2004).

8. Blott, S. J., Croft, D. J., Pye, K., Saye, S. E. & Wilson, H. E. Particle size analysis by laser diffraction. *Geol. Soc. London, Spec. Publ.* **232**, 63–73 (2004).

9. Barber, T. A. *Control of Particulate Matter Contamination in Healthcare Manufacturing. (1st ed.). CRC Press, 1999*. doi:10.1201/9780429246692.

10. Gillespie, S. *et al.* Assessing microbiological water quality in drinking water distribution systems with disinfectant residual using flow cytometry. *Water Res.* **65**, 224–234 (2014).

11. Yang, L. & Yamamoto, T. Quantification of Virus Particles Using Nanopore-Based Resistive-Pulse Sensing Techniques. *Front. Microbiol.* **7**, 1500 (2016).

CHAPTER 2

State-of-the-Art Methods for Particle Detection and Characterisation

1.1   Particles

Particles are mostly interpreted as tiny fragments that constitute all matter. Industrial particles cover a wide range of sizes, from nanometres to hundreds of millimetres. Therefore, the concept of particles is limited within the scope of this thesis. The particle sizes studied span from a few hundred nanometres to approximately a hundred micrometres. **Figure 2.1** shows some common industrial particles and their approximate size ranges.
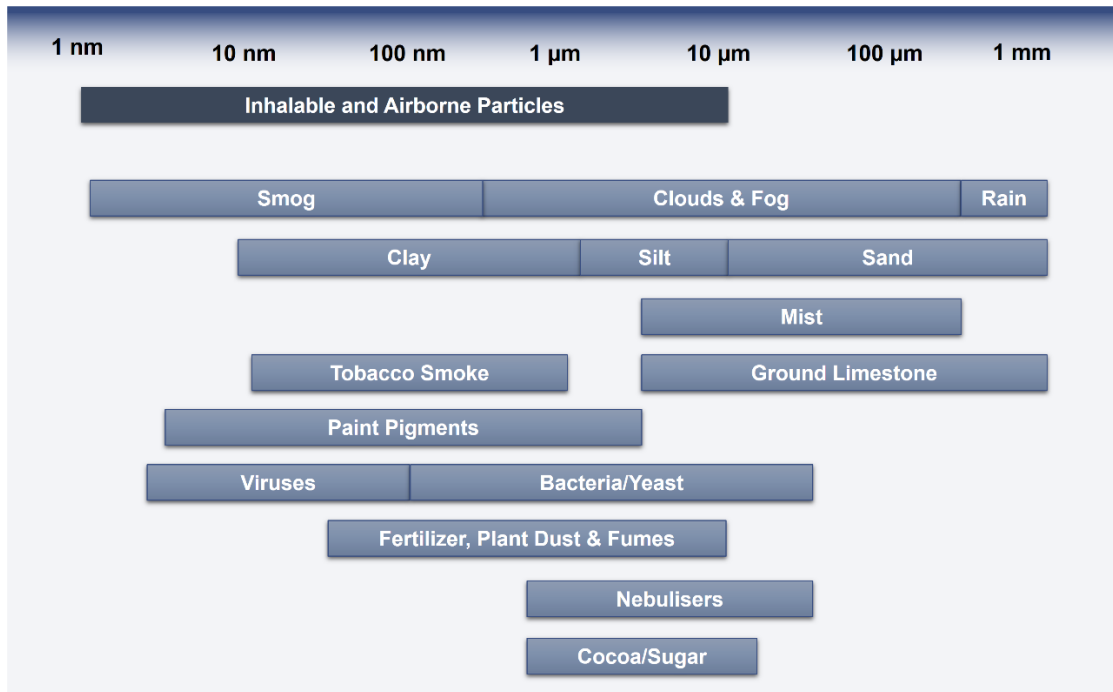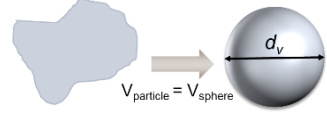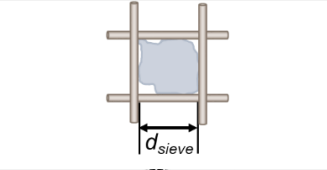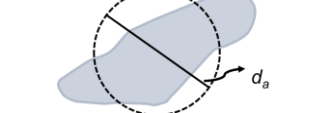
Figure 2.1: Industrial particles.   *Some examples of various particles and their approximate size ranges. Adapted from references[1,2].*

Table 2.1: Equivalent sphere concept for irregularly shaped particles[2,4]

Equivalent Diameter	Symbol	Definition	
Volume diameter	d_v	Diameter of a sphere with the same volume as the particle	d_v $V_{particle} = V_{sphere}$
Surface diameter	d_s	Diameter of a sphere with the same surface area (SA) as the particle	d_s $SA_{particle} = SA_{sphere}$
Stokes' diameter	d_{st}	Diameter of a sphere with the same density, ρ, and settling rate, v, as the particle under conditions defined by Stokes' law	d_{st} $\rho_{particle} = \rho_{sphere}$ $v_{particle} = v_{sphere}$
Sieve diameter	d_{sieve}	Diameter of a sphere equal to the width of the sieve aperture through which the particle will pass	d_{sieve}
Projected area diameter	d_a	Diameter of a circle with the same area as that projected by the particle	d_a

Particles most often present themselves in particulate form, consisting of a large number of particles, as opposed to their bulk form for the same volume or weight. The physical properties of particles in a particulate system can vary, and the ensemble's macroscopically observable behaviour is frequently different from that of the bulk material[1]. Depending on the property being investigated, a particulate system can be classified into monodisperse or polydisperse. All particles in a monodisperse system have the same properties, whereas all or some particles in a polydisperse system have distinct characteristics.

Particle sizing methods can be classified into two categories, namely ensemble and non-ensemble. The particle size distribution (PSD) obtained from the contribution of signals from all particles is acquired using the ensemble approach, which involves measuring all particles at the same time. In non-ensemble approaches, particles are measured one at a time and sorted into distinct bins to determine the PSD, or they are physically separated based on size using an external force or process, and the PSD is derived by measuring the number of the different sizes being separated.

Depending on the method used, there is a variety of commercially available particle sizing devices on the market, each with its own set of benefits and drawbacks. These devices are usually designed to work within a specific particle size range and, hence, are best suited for specific applications. In the following sections, we present a brief overview of the most commonly used traditional methods, as well as some of the advanced optical methods currently employed for particle characterisation in many industries.

2.1 Particle Characterisation Techniques: Non-Optical

2.1.1 Sieving (~ 5 μm – 10 cm)

Sieving[5] is one of the earliest and most straightforward procedures for characterising particles. Fine particles are separated from coarse particles using a set of sieves with varied mesh sizes, as shown in **Figure 2.2**. This is accomplished by vibrating the particles in a combination of vertical, horizontal, and rotational motions, causing the particles to orientate themselves and fall through the mesh apertures into a sieve with a smaller mesh size. After the sieving procedure is completed, the weight of each sieve is measured and compared to the weight of the sieve before adding the sample. This yields the weight of the material in each sieve, which is then used to calculate the PSD and the mean diameter of the sample.
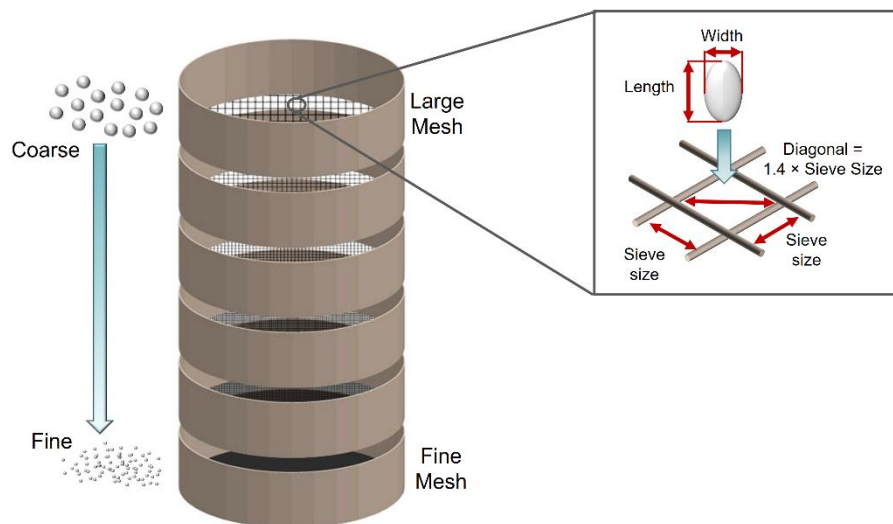
Figure 2.2: Sieve Analysis. Particles are separated by passing them through sieves with different mesh sizes from large to fine [adapted from reference[6]]. The zoomed in portion shows that an elongated particle can pass through the mesh by orientating itself 'on end' [adapted from reference[7]].

Due to the simple working principle of sieving, it is frequently used in many industrial applications and is particularly advantageous in determining particle sizes in the range of several millimetres, as this range is well above the detection limit of other conventional particle sizing methods. However, despite being the most common method, this approach has several drawbacks. Firstly, it ignores particle shape, and a particle's ability to pass through the mesh is determined by its second smallest dimension, i.e., its width. Two elongated particles with comparable widths but different lengths, for example, could pass through the mesh and produce the same result. Hence, this method is unreliable for flat and elongated particles. In addition, it is a low-resolution approach, especially for particles with a narrow size distribution, owing to the limited mesh sizes available commercially. The accuracy of the results is determined by particle characteristics, orientation, mechanics of the sieve movement, and the agitation time. Moreover, sieving is time-consuming, slow, requires a large number of samples for analysis and lacks the ability to provide real-time monitoring of industrial processes.

23

2.1.2 Sedimentation (~ 50 nm – 100 µm)

Sedimentation[8] involves measuring the time taken for particles to settle through a liquid, either due to a gravitational field or through centrifugation, and then relating this to the particle mass using the Stoke's law. The principle behind this technique is the fact that larger particles settle faster than smaller particles in a liquid medium. On determining the settling rate, the particle size can be calculated using its density. Hence, this method requires a prior knowledge of the particle's density, as well as the viscosity and the density of the liquid medium. It is also assumed that the particles are all spherical in shape, have equal densities, and settle in a laminar flow. The concept of equivalent diameter is used when applying this method to non-spherical particles. This means that the calculated Stoke's diameter is equal to the diameter of a spherical particle exhibiting the same properties (geometric, optical, electrical etc.)[9] as the measured particles.

Sedimentation has been widely used in the past due to its low cost, high accuracy and reproducibility, continuous operation and fairly large measurement range. This approach, however, is unsuitable for particles larger than 50 µm due to turbulent-flow (large Reynolds numbers) and smaller than 2 µm because of diffusion due to the Brownian motion. In addition, emulsions (which do not settle), highly dense materials (which settle quickly), and mixes of different densities are not suitable for sedimentation. Also, the viscosity is affected by the surrounding temperature. Moreover, the various types of sedimentation methods that require separation, such as the pipette method, sedimentation balance method, and diver method, are particularly time consuming since fine particles in samples require extensive settling durations (up to 60 minutes). Centrifugal sedimentation with light or X-ray detection is a more recent form of sedimentation and, nowadays, is the most typically used version due to shorter measuring times[10]. Further details on the different forms of sedimentation method can be found in reference[8].

2.1.3 Electrozone Sensing (~ 0.4 µm – 1.2 mm)

Electrozone sensing[11], or the Coulter Counter, is a particle counting method that provides information regarding the particle size, as well as the number and volume distribution. In a Coulter Counter, a tube with an aperture of known diameter is placed in an electrolyte (electrically conducting liquid), and particles (non-conductive) are suspended in it (**Figure 2.3**). The flow of current through the electrolyte is caused by two electrodes, one within the tube and the other outside. When a particle passes through the aperture, i.e., the sensing zone, it displaces a volume of the electrolyte solution equal to its own volume, increasing the resistance between the electrodes. This results in a voltage pulse, the intensity of which is proportional to the volume of the particle and the number of which gives the particle count. The particle size is then derived as the equivalent spherical diameter based on a calibration curve.

The electrozone technique is capable of performing rapid (< 1 min) measurements with high resolution, reproducibility and sensitivity and a minimal sample quantity. The accuracy of the size measurement is usually within 1 to 2 percent. The electrical response of the particles is not influenced by their shape, hence the particle size obtained is unbiased. Other optical properties of the particle, including colour, refractive index, and composition, are likewise unaffected by this approach. The lower size limit of 0.4 µm is determined by the signal-to-noise ratio and the upper size limit (~1200 µm) is set by the ability to suspend particles uniformly in the electrolyte. Two or more apertures must be used for measurements across a broad range, with the data being combined to produce a complete size distribution. The main drawbacks of this approach are that it cannot measure conductive particles like metals and that it necessitates the use of an electrolyte solution. Moreover, this method cannot produce reliable data for porous samples because the

pores can become filled with the electrolyte solution and the effective volume displaced may be much smaller than its true "envelope volume"[12].
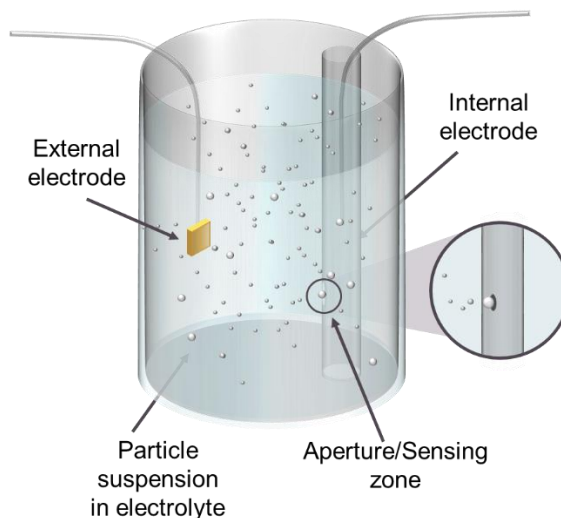
Figure 2.3: Working principle of electrozone sensing or Coulter Counter.

While the three methods explained above are the most common non-optical techniques, there are also other, less popular, particle characterisation techniques. These include chromatographic methods and acoustic spectroscopy for sub-micron particle sizing, as well as some other more application-specific methods, for example, electroacoustic spectral analysis for zeta potential determination and particle sizing, pulsed field gradient nuclear magnetic resonance (PFG-NMR) for determining diffusion coefficient, and many more. An elaborate description of each of these methods with their benefits and drawbacks can be found in references[1,13].

2.2 Particle Characterisation Techniques: Optical

When a light beam illuminates a particle, a part of the incident light is either scattered or absorbed or both depending on the wavelength of the incident light, and the physical (complex refractive index) and geometrical (size and shape) properties of the particle. This combined effect results in the extinction of the incident light. The majority of optical characterisation techniques rely on determining the relationship between the characteristics of the target particles and such properties as the intensity, phase, etc. of the light scattered or absorbed by them. The most widely used methods for particle characterisation are based on light scattering and, hence, we will limit our discussion to only these techniques in the following sections. A brief theoretical background is provided on light scattering to better understand the concept behind these methods.

2.2.1 Light Scattering Phenomena

When an electromagnetic (EM) wave (i.e., a light beam) illuminates a particle, the electrons in the particle oscillate around the positive nucleus. Because of this charge separation, the particle acquires a dipole moment, which oscillates at the frequency of the incident wave, reradiating secondary EM waves in all directions. These secondary waves are referred to as scattered waves, and what we see from a particle is a superposition of the incident wave from the source and the scattered waves from the particle or scattering centres. In an isotropic and homogeneous material,

the scattered waves from individual particles causes destructive superposition and, therefore, no scattering occurs[1].

Most particle sizing techniques use elastic scattering of light to determine the size of various particles. Elastic scattering occurs when the frequency (or wavelength) of the scattered wave is the same as the incident wave. Examples of this type of scattering include Rayleigh and Mie scattering. Mie scattering is applied essentially for spherical particles with diameters, D, equal to or greater than the incident wavelength λ, in contrast to Rayleigh scattering, which is used for particles much smaller than the wavelength of the incident light (**Figure 2.4**).
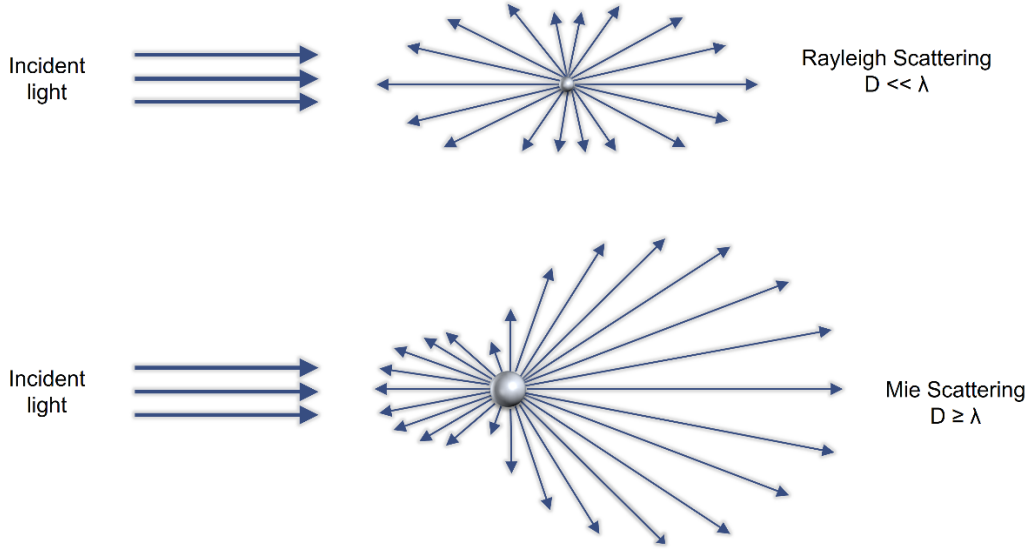
Figure 2.4: Rayleigh and Mie scattering. The scattering pattern is dependent on the particle diameter D and wavelength of the incident light λ.

When characterising particles using light scattering, most often the Mie theory is used due to the high accuracy this model can provide in analysing samples with a wide size range and varied optical properties. The Mie theory, developed by Gustav Mie in 1908, gives the analytical solution to Maxwell's classic equations for the scattering of light by a single homogeneous, isotropic and spherical particle of arbitrary diameter immersed in a non-absorbing medium with a refractive index that is different from the particle itself. In brief, Maxwell's electromagnetic equations in spherical co-ordinates are solved by applying the standard boundary conditions at the surface of the particle. To match the inside and outside electromagnetic fields at the boundary of the spherical particle, the incoming plane wave is expanded into vector spherical harmonics using Legendre polynomials. For practical applications, the solution sought is in the far-field zone, i.e., at a distance, *r*, much larger than the incident wavelength.

In the far-field region, the Mie theory gives the relationship between the incident and scattered components of the electromagnetic field that are parallel and perpendicular to the scattering plane. Since practical scattering experiments measure intensity, in general, for unpolarised light, the total scattered intensity by a single spherical particle can be expressed as[14]:

$$I(\theta) = \frac{I_0(|S_1(\theta)|^2 + |S_2(\theta)|^2)}{2k^2r^2} \tag{2.1}$$

where, wavenumber $k = \frac{2\pi}{\lambda}$, with λ the wavelength of incident light, $I(\theta)$ is the total scattered intensity, I_0 is the incident light intensity, $S_1(\theta)$ and $S_2(\theta)$ are the scattering amplitude functions with θ being the scattering angle with respect to forward direction and are given by the following equations[14]:

$$S_1(\theta) = \sum_n \frac{2n+1}{n(n+1)} (a_n \tau_n + b_n \tau_n), \qquad (2.2)$$

$$S_2(\theta) = \sum_n \frac{2n+1}{n(n+1)} (a_n \tau_n + b_n \pi_n), \qquad (2.3)$$

In the above equations, τ_n and π_n are the Mie angular functions and a_n and b_n are the Mie coefficients for the scattered field given by[14]:

$$\pi_n = \frac{P_n^1}{sin\theta} \qquad (2.4)$$

$$\tau_n = \frac{dP_n^1}{d\theta} \qquad (2.5)$$

$$a_n = \frac{m\psi_n(mx)\psi'_n(x) - \psi_n(x)\psi'_n(mx)}{m\psi_n(mx)\xi'_n(x) - \xi_n(x)\psi'_n(mx)} \qquad (2.6)$$

$$b_n = \frac{\psi_n(mx)\psi'_n(x) - m\psi_n(x)\psi'_n(mx)}{\psi_n(mx)\xi'_n(x) - m\xi_n(x)\psi'_n(mx)} \qquad (2.7)$$

where, P_n^1 is the associated Legendre function of first kind of degree n, $x = ka$ is the size parameter, a is the radius of the particle, $m = \frac{n_p}{n}$ is the relative refractive index between the particle (n_p) and the medium (n), ψ_n and ξ_n are related to the *Ricatti-Bessel* functions. The detailed derivation of the Mie theory can be found in reference [14].

In order to better understand how the scattered intensity varies with the scattering angle for spherical particles of various sizes, polar plots of different particle diameters are shown in **Figure 2.5**. The size parameter has a major influence on the angular distribution of scattered light from spherical particles. Large particles have a narrow scattering lobe that is predominantly in the forward direction, and the scattering intensity is substantially higher than that of smaller particles at the same scattering angle. The angular dependence of scattered light diminishes as the particle size decreases, becoming practically flat for particles smaller than 50 nm in diameter. Based on these characteristics, the size of particles can be derived using light scattering measurements.
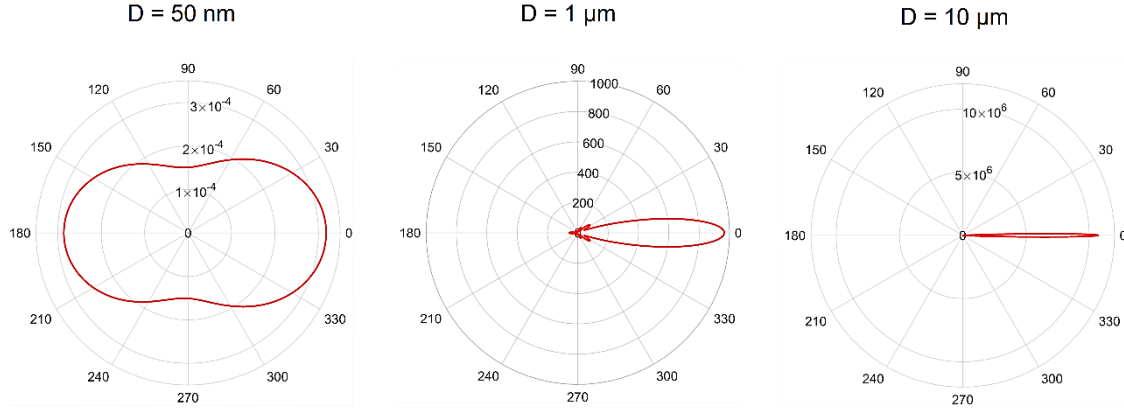
Figure 2.5: Mie scattering patterns from spherical particles of different diameters. *Polar plot of particles (n_p =1.59) with diameter 50 nm (left), 1 µm (middle) and 10 µm (right) suspended in water (n =1.33), showing how the angular distribution of scattered intensity varies with particle size. The graphs are plotted using BHMIE[14] code in Matlab.*

With advancements in computation techniques, calculation of Mie scattering for a single sphere has become relatively easier. Reference[15] provides a compilation of several computer source codes based on Mie theory. Among those, the BHMIE code by Bohren and Huffman[14] is the standard algorithm and is used in this thesis for most particle scattering simulations.

In some particle size analysers, Fraunhofer diffraction theory[14,16], a simplified model of Mie theory, is sometimes used especially for medium or coarse particles which are much larger than the wavelength of light (typically > 25 µm). In Fraunhofer theory, diffraction or scattering of light is considered only at the contour of the particle. This theory assumes particles to be opaque discs which means that the scattering pattern obtained from these discs will be the same as that obtained from circular apertures[2] according to Babinet's principle. The diffraction pattern is usually measured at a large distance ("far-field") from the scatterer and since this theory applies to large particles the scattering intensity is concentrated mostly in the near forward direction, i.e. at angles smaller than 10 degrees. For small scattering angles θ, the intensity distribution according to Fraunhofer diffraction theory can be expressed by the following equation[14,16]:

$$I(\theta) = \frac{I_0}{k^2 r^2} x^4 \left[\frac{J_{1(x\theta)}}{x\theta}\right]^2 \tag{2.8}$$

where J_1 is the Bessel function of the first kind of order unity.

The main advantage of the Fraunhofer theory is that it is relatively straightforward and does not require any prior knowledge of the optical properties (e.g., refractive index) of the samples to be measured. However, this method can only be used for larger particles and cannot provide the precision offered by the Mie theory, especially for small particles.

2.2.2 Laser diffraction (~ 10 nm – 4 mm)

Laser diffraction (LD)[17–19] or Static Light Scattering (SLS) is the most widely used technique for particle size analysis today. In this particle characterisation method, the particle size is derived by directly probing the angular distribution of scattered light from the sample, and, unlike other methods, once the device is properly set, this method does not require any prior calibration or scaling. Despite the fact that most modern equipment employs the generalised Mie theory to determine particle size, this method is still known as *laser diffraction*. This is because laser

28

diffraction devices were previously constrained to employing solely the Fraunhofer diffraction theory for particle size calculation due to limited computation power.

In a classical LD setup, as shown in the illustration in **Figure 2.6**, a laser is used as the light source, followed by a beam processing unit that includes a beam expander and other components such as lenses and pinholes to produce a collimated beam of light. Particles dispersed in a liquid or air (in a sample chamber) are illuminated by the laser beam and the angular distribution of scattered light from the particles is detected by a multi-element photodetector array placed mostly in the near forward direction. The unscattered forward light is also collected by the detector element, and can later be used for determining the light obscuration of the sample. The optical signal is then converted into an electrical signal by the detector array, and the particle size distribution is determined using computer software based on an appropriate scattering model.

The detector geometry is determined by the desired angular range, angular resolution, and sensitivity, among other factors, and can be of various types, such as concentric half-rings, X patterns, fly-wing patterns, and so on[1]. Most often ring detectors are used, i.e., the detector elements are arranged in a logarithmic fashion, with the size (radii and width) of each detector increasing as a constant multiple of the size of the preceding detector element. Since the total area of each detector element increases with an increasing angle, this arrangement aids in improving the signal-to-noise ratio when measuring small particles that scatter light at a low intensity.

The volume equivalent sphere diameter is used to report particle size in LD devices. Three parameters are frequently used to define the width of a sample's particle size distribution. The D50 which is the median diameter dividing the sample population into equal halves, the diameter D10 meaning that 10 percent of the sample diameter falls below this value, and the D90 is the diameter below which 90 percent of the sample diameter lies. Since the angular scattering profile recorded by the detector array is a superposition of various intensity patterns generated by individual particles of various sizes and shapes, LD is referred to as an ensemble method.
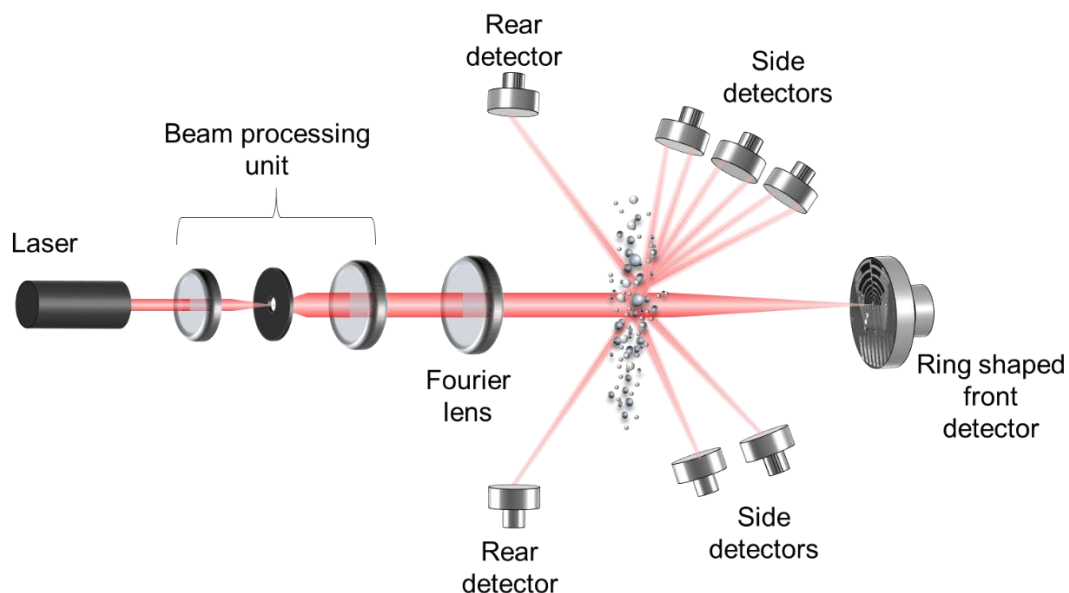
Figure 2.6: A typical LD setup used in commercial particle size analysers. *When particles are irradiated by a laser beam, a scattering pattern is created, which is observed by front, side and rear detectors.*

The measuring range of LD devices is mostly determined by the distance between the laser, sample chamber and photodetector, as well as the position of and distance between the detector

elements. The typical measurement size range is from 0.1 µm to more than 1 mm, although new instruments can now perform measurements down to the nanometre level. The upper size limit is set by the fact that large particles scatter at small angles, making it difficult to resolve particles of similar sizes. The lower limit is defined by the weak scattering intensity from small particles, which can be enhanced by using a shorter incident wavelength such as blue or violet. Thus, many current instruments employ multiple laser sources for improved performance over a broad particle size range.

There are two types of optical setups commonly used in recent commercial laser diffraction system designs, namely the Fourier optics or reverse Fourier optics setups[20]. As shown in **Figure 2.7**, in a traditional Fourier optics configuration the collection lens is placed after the measuring zone, which has the advantage that regardless of the position of a particle in the laser beam, scattered light at a specified angle will be refracted by the lens to fall at a specific detector. This is especially useful for spray measurements since the particles can be dispersed over a larger measuring zone. Because the scattering angle is proportional to the particle size (small particles scatter at large angles), the measuring range in this configuration is mainly determined by the focal length of the lens.
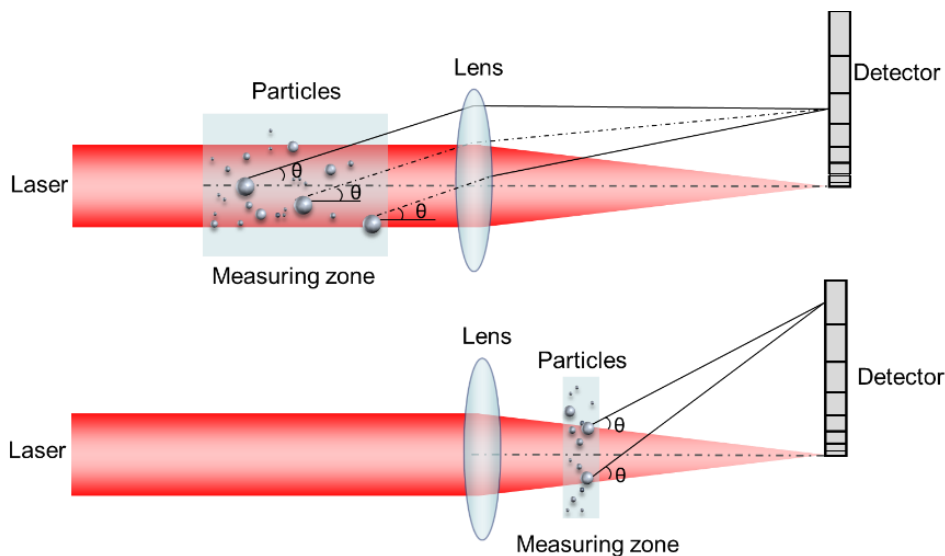
Figure 2.7: Lens arrangements for laser diffraction systems. Fourier optics setup (top) and reverse Fourier optics setup (bottom). Adapted from reference[21].

In the reverse Fourier setup, the collection lens is placed before the measuring zone and, thus, particles are illuminated by a convergent beam. The distance between the measuring zone and the detector determines the measurement range in this arrangement, and depending on the location of the particle in the converging beam, the scattered light at different scattering angles will arrive at a particular detector. Unfortunately, this results in a blurred diffraction pattern compared to the Fourier setup, leading to sizing errors. Therefore, in order to improve particle sizing accuracy, the sample chamber must be narrow in this setup. The advantage of the inverse Fourier method is that it allows for the collection of wider scattering angles as additional detectors can be placed around the sample chamber. Hence, this setup provides a broader measurement range with an improved resolution.

The data evaluation procedure in LD devices is based on the solution to the following integral equation[16], which is the sum of the scattering intensities from an ensemble of particles:

$$I(\theta) = \int A(\theta, x) f(x) dx$$

(2.9)

where, x is the size parameter, the kernel $A(\theta, x)$ is the scattered light intensity at angle θ of a single spherical particle with size parameter x, and $f(x)$ is the particle size distribution. This is the so-called inverse problem, the numerical solution of which is complicated and, typically, an ill-posed problem[1,22]. Many inversion methods have been developed over the years, details of which can be found in literature references[22–25]. In practice, the size distribution $f(x)$ is discretised into f sizes and equation (2.9) is reduced to a series of linear algebraic equations that can be represented by the following matrix equation:

$$I = A \cdot f$$

(2.10)

where,

$$I = \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_m \end{bmatrix} \qquad f = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}$$

(2.11)

$$A = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}$$

(2.12)

In equation (2.11), I_i $(i=1, 2 ... m)$ is the scattering intensity detected by each element of the ring detector, i.e., at each detection angle. The measurement target particle size range is divided into size classes where each class is a group of similar sized particles defined by two diameters. Therefore, each element of the particle size distribution f, f_j $(j=1, 2 ... n)$ represents the percentage of particles in the size class (x_j, x_{j+1}). The matrix A in equation (2.12) contains the scattering coefficients calculated using the Mie theory. Each element of A, a_{ij} $(i=1, 2 ... m,$ and $j=1, 2...n)$ represents the expected scattering intensity at each detection angle for each size class. The complex refractive index, i.e. both the real and the imaginary part, of the particles and medium must be chosen beforehand in order to calculate the elements of A using the Mie scattering theory. The PSD is then determined by inversion of equation (2.10):

$$f = A^{-1} \cdot I$$

(2.13)

where, A^{-1} is the inverted matrix of equation (2.12). In commercial instruments, this matrix inversion is carried out by proprietary algorithms.

LD has become the state-of-the-art technology for particle size determination due to the wide size range that can be measured using this method. Commercial LD devices are fast and robust, capable of providing results with high accuracy and reproducibility. However, there are limitations to this method, including the need to optimise the sample concentration in order to achieve a reasonable signal-to-noise ratio while, at the same time, ensuring elastic single particle scattering. Also, when measuring concentrated samples, multiple scattering becomes a crucial problem, but, using different approaches, several analytical correction factors[26] have been developed in recent years. Nowadays, commercial LD devices are required to introduce these complex correction factors into their proprietary algorithms in order to reduce the effect of multiple scattering and retrieve the PSD with greater accuracy. Another drawback of this method is that it requires previous knowledge of the refractive index of the sample as well as the medium in which the particles are dispersed. This demands a trial-and-error process as incorrect refractive indices used for establishing the optical model can result in inaccurate size measurements. Moreover, this method assumes that all particles measured are perfect spheres. It has been

demonstrated that the size of non-spherical particles are underestimated by LD[27], thus reducing the measurement accuracy. Furthermore, commercial LD devices are large and expensive products, limiting their use in online industrial applications in most cases.

2.2.3 Dynamic Light Scattering (~ 0.1 nm – 10 μm)

For submicron particle measurement, dynamic light scattering (DLS)[1,28] has now become an industry standard technique. The optical setup consists of a laser source that illuminates particles in suspension, and the fluctuations of the scattered light intensity are recorded over a certain time period by a fast photon detector, such as a photodiode or photomultiplier tube (PMT), placed at a specific scattering angle (**Figure 2.8**). Using a single detector limits the particle size range that can be measured so, for polydisperse samples, detectors placed at several angles are required to determine the particle size distribution accurately. Depending on the sample concentration, detectors can be placed at 90° (side scattering) or 175° (back scattering).
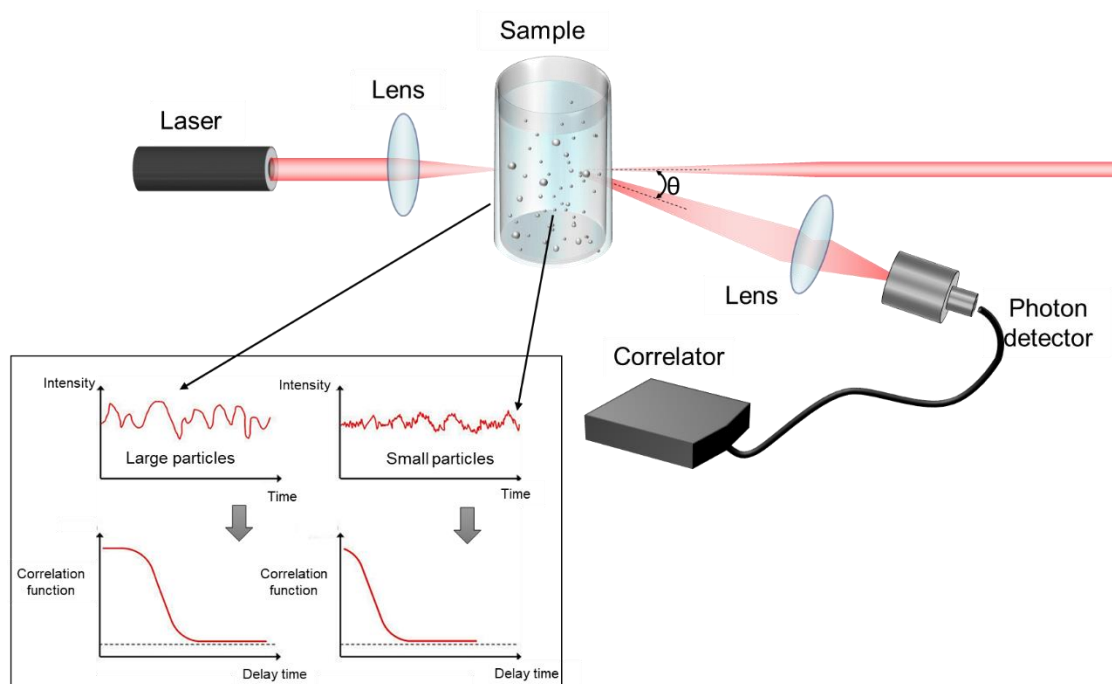
Figure 2.8: A typical setup of a DLS system. *Examples of intensity fluctuations for two types of particle sizes and their corresponding autocorrelation functions are also shown. Adapted from references[29,30].*

Due to the random Brownian motion of the particles, the optical signal acquired from the detector shows fluctuations in signal intensity. This signal is then analysed in real-time by a digital correlator to determine the autocorrelation function, which essentially indicates how long a particle stays at a specific location in the sample. As smaller particles move faster, the autocorrelation function exhibits a steeper exponential decay compared to that obtained with large particles. The diffusion coefficient is then derived from the autocorrelation function, and with a prior knowledge of the refractive index of the liquid, the wavelength of the incident light and the scattering angle, the hydrodynamic size of the particles is calculated using the Stokes-Einstein relationship given below:

$$D_h = \frac{k_B T}{3\pi\eta D_t}$$

<div align="right">(2.14)</div>

where, D_h is the hydrodynamic particle diameter, k_B is the Boltzmann's constant, T the thermodynamic temperature, η the dynamic viscosity and D_t is the translational diffusion coefficient.

The main advantages of the DLS method are that it is non-invasive in nature, requires only a small sample size and does not require complex sample preparation steps. This method is suitable for measuring particles in the lower nanometre range and up to a few micrometres. The lower size limit of a few nanometres is usually determined by the signal-to-noise ratio (SNR), where the scattered intensity from the particles must be greater than the noise associated with the measurement system. The upper limit depends on the sedimentation effect, which means there is no random motion from the particles, leading to inaccuracy in the measurements. The particle concentration in DLS measurements should ideally be low enough to avoid multiple scattering effects while still achieving a good SNR, which can be difficult, especially when measuring food particles. Dilution can create changes in structural variation, leading to unreliable results. However, in state-of-the-art devices, Non-Invasive Back-scatter[31] technology is employed to increase the concentration range of the particles to be measured. Another limitation is the need to maintain a constant temperature and viscosity, as DLS measurements are highly sensitive to any changes in these parameters. DLS is a low-resolution method and the presence of large particles or aggregates can also cause measurement inaccuracies[28] since the scattering intensity is nearly proportional to the 6th power of the particle diameter. Thus, prior to each measurement, it is important to clean the sample cell properly in order to achieve reliable results.

Other scattering approaches for detecting nanometre-sized particles have also recently emerged, such as nanoparticle tracking analysis (NTA)[32], which uses image recording to monitor individual particle movement via scattering. NTA, like DLS, uses the diffusion coefficient to determine the hydrodynamic size of the particles, but it is capable of overcoming some of the limitations of DLS [32,33].

2.2.4 Flow-Cytometry (~ 0.2 μm – 150 μm)

Flow-cytometry[34,35] is a powerful technology for the rapid detection of particles, most often cells, using light scattering and fluorescence. This technique is most widely used to characterise and identify distinct types of cells in a heterogeneous population, to quantify cells based on their optical properties, and to sort cells into different groups for further evaluation. Flow-cytometry is a versatile technology that allows multiple cell parameters to be analysed simultaneously at the single-cell level.

In a traditional flow-cytometer (**Figure 2.9**), single cell analysis is accomplished via hydrodynamic focusing. In other words, the sample consisting of a cell suspension is injected into a faster-moving sheath fluid, usually a saline solution, inside a flow chamber. This flow chamber is considered the heart of the flow-cytometer since its design causes the sample to be focused into a 'sample core' at the centre of the sheath fluid. A laminar flow is achieved within the flow-cell due to the pressure differences between the two fluids, and the sample core is maintained in the centre of the sheath fluid. This forces the cells or particles in suspension to pass through the interrogation point (laser intercept) in a single file, where they interact with a focused laser beam.
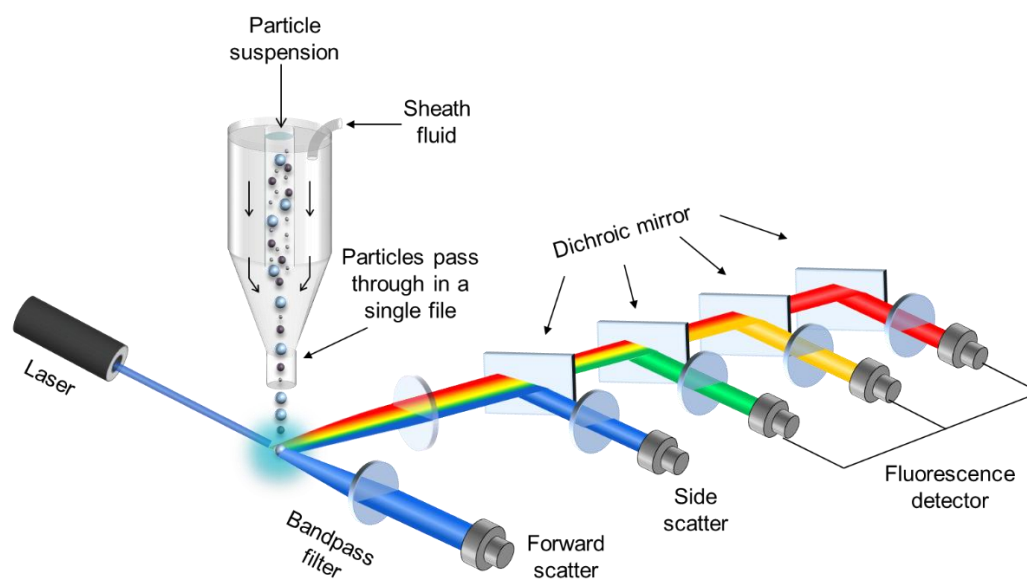
Figure 2.9: Schematic of a conventional flow-cytometer setup. *Adapted from reference[36].*

The light scattered by the cells is measured by a photodiode or a PMT, either in the forward direction (forward scatter, FSC), which provides the relative particle size, or at an angle of 90° (single scatter, SSC), which gives information regarding the structural complexity or granularity of the cell. Fluorescence measurements can also be performed, simultaneously, if the sample is previously stained with fluorescent dyes or with fluorescently tagged antibodies. As the fluorescing cells or particles are excited by the laser beam, they emit fluorescence at a longer wavelength than the incident light. Using several dichroic mirrors and emission filters, the fluorescent and scattered signal pulses from individual cells are directed to the appropriate detectors, where they are converted and amplified to voltage pulses often referred to as 'events'.

Nowadays, flow-cytometers can measure and analyse up to 30 parameters simultaneously (28 for fluorescence measurements and one each for FSC and SSC measurements), thanks to advancements in hardware technology, data analysis capabilities, and new fluorescent probes. Depending on the flow rate, the analysis can be carried out at a high speed (up to 200 k events/s). Additionally, by replacing PMTs with avalanche photodiodes (APD) in numerous flow cytometers, enhanced sensitivity for fluorescence detection has been achieved. These recent developments have allowed the widespread application of flow-cytometry in many different fields from virology, cancer biology and immunology to monitoring infectious diseases and many more. However, conventional flow-cytometers are high-maintenance and complicated instruments that require highly skilled personnel to operate them. Furthermore, these devices are large and expensive, costing tens of thousands of euros, in part due to the hydrodynamic focusing required for one-by-one cell analysis. These limitations make them unsuitable for point-of-care (POC) diagnostics or for developing countries with limited resources.

In recent years, different variations of flow-cytometry instrumentation have evolved, combining traditional flow-cytometry with other methods, such as fluorescence microscopy in imaging cytometers or time-of-flight mass spectrometry in mass cytometers. A detailed description of the different types of flow-cytometers can be found in reference[35].

2.2.5 Other Methods: Light Obscuration

Single Particle Optical Sensing

In addition to light scattering methods, light obscuration, or single particle optical sensing (SPOS), can also be used to detect particle size and concentration in liquid suspension. This method, unlike ensemble methods such as LD, uses a diluted sample suspension to ensure that particles arrive in the sensing zone one at a time. When a particle flows within this sensing zone, it interacts with the laser beam, leading to light extinction, which is then detected by a photodetector. The amount of light that is obscured due to the presence of the particle is proportional to its size. The photodetector converts the optical signal into a voltage pulse, from which the individual particle's size is calculated by comparing the pulse height to a calibration curve previously obtained from measurements of standard particles with known diameters. The measurement range of most commercial instruments is from 2 µm to 1000 µm approximately. Below 1 µm, light scattering detection is required, which is achieved by placing a photodetector at an oblique angle, enabling the detection limit to be lowered to 0.5 µm. In addition to scattering measurements, some devices use a focused beam from a high-power laser to further lower the detection limit to 0.15 µm.

SPOS is a high-resolution, high-throughput technology with a wide dynamic range and excellent accuracy. This approach can be used to detect contamination (e.g., clean room environment, water, hydraulic fluids, etc.), as well as to analyse samples such as emulsions, injectables, proteins, and many others. However, sample concentration is a critical issue when employing this method, with most commercial devices either using an automatic dilution procedure or requiring manual predilution to ensure that a sufficiently diluted sample is used for measurement.

Image Analysis

Image analysis is another commonly used particle characterisation technique worth mentioning here. Manual microscopy, or static image analysis, is a traditional method for determining the morphological information of particles. However, despite the high resolution and great image quality that can be achieved, this process is not only time-consuming but also labour-intensive. Furthermore, this technique is best suited to analysing samples with narrow size distributions. Over the past few decades, advances in camera technology and data processing software have facilitated the development of dynamic image analysis[37] (DIA) systems. These devices can analyse over 300 images per second, in real-time, of particle suspensions in flow and can rapidly generate a volume or number distribution of the entire sample population. Depending on the application, different size definitions, such as chord dimension, Martin diameter[2] or Feret diameter[2], are used to determine the particle size. Circularity, convexity, aspect-ratio, symmetry, and other shape properties can also be examined with DIA. The field of view of the camera determines the upper size limit that can be assessed with this technology, while the resolution determines the lower limit, with a typical measuring range being around 0.8 µm to 135 mm. DIA is a high-resolution technology that enables rapid measurements with great accuracy and repeatability across a wide range of sizes. The reliability and repeatability of the measurements are determined by the sample volume measured. For accurate results, polydisperse samples, for example, require a considerable sample size to be analysed. Some factors[38] that can contribute to uncertainties in DIA readings are motion blurring, particle overlapping, the optical system's depth of focus, and particle orientation.

A summarised table comparing the different particle characterisation methods discussed in sections 2.1 and 2.2 is provided in **Table 2.2.**

Table 2.2: Comparison of various non-optical and optical particle characterisation techniques

Particle characterisation technique	Size	Shape	Measurement range	Resolution	Measurement time	Sample types
Sieving	Yes	No	5 µm – 10 cm	10 – 40% of size depending on sieve series chosen	5-30 min; 1-2 hr for wet sieving	Dry powders (dry or wet sieving); suspensions (wet sieving)
Sedimentation	Yes	No	50 nm – 100 µm	3 –10% depending on the type of sedimentation	0.25-8 hr depending on PSD and required resolution	Dispersed dry powders, suspensions, emulsions
Electrozone sensing	Yes	No	0.4 µm – 1.2 mm	3% relative	typically 1-5 min	Particle suspensions in conductive liquid
Laser diffraction	Yes	No	10 nm – 4 mm	10 – 40% relative, depending on particles and size of detector elements	< 1 min	Dispersed dry powders, sprays, suspensions, emulsions
Dynamic light scattering	Yes	No	0.1 nm – 10 µm	30% relative	about 1 min	Suspensions/ emulsions
Flow cytometry	Yes	No	0.2 µm – 150 µm	< 5% approximately	depends on the flow rate and the measured volume	Particle suspension
Single particle optical sensing	Yes	No	2 µm – 1000 µm	Typically < 5%	< 10 s	Particle suspension
Image analysis	Yes	Yes	0.8 µm – 135 mm	≥ 0.5% relative	about 1– 60 min	Dispersed dry powders, suspensions, emulsions

2.3 Commercial Optical Particle Detection Devices

The discussion in the previous section illustrates that optical detection techniques invariably provide far greater benefits compared to other technologies available for particle characterisation. For particle size analysis, it is evident that devices based on light scattering are capable of measuring a wide range of sizes, and from the various light scattering-based methods, LD provides excellent performance across a wide measurement range. Different particle characterisation methods have also evolved into online versions, offering a variety of solutions for different applications. Particle size range, shape analysis and integrability constraints in each application dictate the recommended method to use.

Currently, there are several providers offering commercial systems for various applications in industrial environments, including Malvern, Horiba, Beckman Coulter and Shimadzu (). The Malvern *Insitec* series, for example, is an LD-based online analyser with a working size range

from 0.01 µm to 3500 µm, capable of continuous, real-time particle size analysis for monitoring industrial processes. On the other hand, the Malvern *Mastersizer* range is one of the most popular offline, benchtop analysers for measuring particles from 0.1 µm to 2500 µm. Horiba also has a similar line of particle size analysers, together with particle counters, based on light scattering for applications in the automotive industry. For clinical diagnostics, Beckman Coulter provides a range of flow-cytometry instruments for specific applications. **Table 2.3** provides a comparison between different commercial instruments based on some critical specifications.

Figure 2.10: Commercial particle characterisation instruments. *From left, Malvern's Insitec Dry[a] and Mastersizer 300[b], Horiba's particle counter MEXA-2000SPCS[c] (right) and Beckman Coulter's flow-cytometer CytoFLEX S[d].*

However, despite their robustness and ability to provide rapid and accurate measurements, these devices rely on large-scale instrumentation that has a high maintenance cost and requires skilled professionals. To address these issues, the particle detection and characterisation field is currently focusing on developing miniaturised, rapid, portable, easy-to-handle and cost-effective systems. Such instruments will allow not only large industries but also small and medium scale manufacturers to install several platforms in their facilities to monitor certain processes both offline and online. Compact particle characterisation systems are also appealing in the health sector, particularly for point-of-care applications in low-resource settings.

Specification	Malvern, Mastersizer 3000	Malvern, Insitec dry	Horiba, Partica LA-960V2	Horiba, MEXA-2000SPCS	Beckman Coulter, CytoFLEX S
Measurement principle	Laser diffraction	Laser diffraction	Laser diffraction	Laser scattering particle counting	Laser scattering and fluorescence
Particle size range (µm)	0.01 – 3500	0.1 – 2500	0.01 - 5000	Lower size limit ~ 23 nm with 50%	Lower size limit ~ 60 nm

[a] https://www.labbulletin.com/articles/Big-jump-in-demand-for-Malverns-real-time-process-control-consultancy/

[b] https://www.agfa.com/agfa-labs/news-events/new-apparatus-malvern-mastersizer-3000-particle-size-analyzer/

[c] https://www.horiba.com/en_en/automotive/products/detail/action/show/Product/mexa-2000spcs-series-57/

[d] https://www.beckman.de/en/flow-cytometry/instruments/cytoflex-s

				counting efficiency	
Dimensions (W, D, H in cm)	$69 \times 30 \times 45$	$66 \times 97 \times 160$	$70.5 \times 56.5 \times 50$	$43.4 \times 73.1 \times 63.7$	$42.5 \times 42.5 \times 34$
Weight (kg)	30	25	54	115	23.4
Price (k€)	$100 - 200$	> 150	~ 85	> 150	~200
Typical measurement time (sec)	< 10	Continuous	~ 60	Continuous	30,000 events/sec

Table 2.3: Overview of providers and latest light scattering based particle characterisation instruments on the market with some key specifications.

In the following chapter, we will introduce the concept of a novel PSA based on light scattering, employing a consumer electronic camera and machine learning. The proposed PSA technology is able to overcome the problems of low integrability and high cost related to current commercial analysers. Due to its small size, this newly designed PSA has significant potential for use outside a standard laboratory, for example, in online and inline industrial process monitoring.

2.4 References

1. Xu, R. Light scattering: A review of particle characterization applications. *Particuology* vol. 18 11–21 (2015).

2. Merkus, H. G. Particle Size, Size Distributions and Shape. in *Particle Size Measurements* 13–42 (Springer Netherlands, 2009). doi:10.1007/978-1-4020-9016-5_2.

3. Jennings, B. R. & Parslow, K. Particle size measurement: the equivalent spherical diameter. *Proc. R. Soc. London. A. Math. Phys. Sci.* **419**, 137–149 (1988).

4. Staniforth, J. N. & Taylor, K. M. G. Particle size analysis | Basicmedical Key. https://basicmedicalkey.com/particle-size-analysis/ (2016).

5. Allen, T. Particle size analysis by sieving. in *Powder Sampling and Particle Size Determination* 208–250 (Elsevier, 2003). doi:10.1016/B978-044451564-3/50006-1.

6. Sieve Analysis - Particle Technology Labs. https://www.particletechlabs.com/analytical-testing/particle-size-distribution-analyses/sieve-analysis (2021).

7. Sifting Through Sieve Analysis | Innopharma Technology. https://www.innopharmatechnology.com/news/sifting-through-sieve-analysis.

8. Allen, T. *Particle Size Measurement*. (Springer US, 1981). doi:10.1007/978-1-4899-3063-7.

9. Calvert, J. G. Glossary of atmospheric chemistry terms (Recommendations 1990). *Pure Appl. Chem.* **62**, 2167–2219 (1990).

10. Allen, T. A Review of Sedimentation Methods of Particle Size. in *Particle Size Analysis* 454 (Royal Society of Chemistry, 2007).

11. Figueiredo, M. M. Electrozone Sensing in Particle Size Analysis. in *Encyclopedia of*

Analytical Chemistry (John Wiley & Sons, Ltd, 2006). doi:10.1002/9780470027318.a1504.

12. Horák, D., Peška, J., Švec, F. & Štamberg, J. The influence of porosity of discrete particles upon their apparent dimensions as measured by the Coulter principle. *Powder Technol.* **31**, 263–267 (1982).

13. Allen, T. *Powder Sampling and Particle Size Determination*. *Powder Sampling and Particle Size Determination* (Elsevier, 2003). doi:10.1016/B978-0-444-51564-3.X5000-1.

14. Bohren, C. F. & Huffman, D. R. *Absorbing and scattering of light by small particles*. *Wiley Interscience* (1983). doi:10.1002/9783527618156.

15. Mie - scatterlib. http://scatterlib.wikidot.com/mie.

16. Hulst, H. C. van de. Light scattering by small particles. By H. C. van de Hulst. New York (John Wiley and Sons), London (Chapman and Hall), 1957. Pp. xiii, 470; 103 Figs.; 46 Tables. 96s. *Q. J. R. Meteorol. Soc.* **84**, 198–199 (1958).

17. Black, D. L., McQuay, M. Q. & Bonin, M. P. Laser-based techniques for particle-size measurement: A review of sizing methods and their industrial applications. *Prog. Energy Combust. Sci.* **22**, 267–306 (1996).

18. Ma, Z., Merkus, H. G., De Smet, J. G. A. E., Heffels, C. & Scarlett, B. New developments in particle characterization by laser diffraction: size and shape. *Powder Technol.* **111**, 66–78 (2000).

19. Blott, S. J., Croft, D. J., Pye, K., Saye, S. E. & Wilson, H. E. Particle size analysis by laser diffraction. *Geol. Soc. London, Spec. Publ.* **232**, 63–73 (2004).

20. 13320. Particle size analysis-laser diffraction methods. *ISO Stand. Auth.* (2009).

21. Laser Diffraction Technology for Routine Particle Size Analysis. 1–6 https://www.azom.com/article.aspx?ArticleID=11363 (2015).

22. Ye, M. *et al.* Inversion of particle-size distribution from angular light-scattering data with genetic algorithms. *Appl. Opt.* **38**, 2677 (1999).

23. Twomey, S. On the Numerical Solution of Fredholm Integral Equations of the First Kind by the Inversion of the Linear System Produced by Quadrature. *J. ACM* **10**, 97–101 (1963).

24. Vargas-Ubera, J., Aguilar, J. F. & Gale, D. M. Reconstruction of particle-size distributions from light-scattering patterns using three inversion methods. *Appl. Opt. Vol. 46, Issue 1, pp. 124-132* **46**, 124–132 (2007).

25. Liu, J. Essential parameters in particle sizing by integral transform inversions. *Appl. Opt.* **36**, 5535 (1997).

26. Jones, A. R. Light scattering for particle characterization. *Prog. Energy Combust. Sci.* **25**, 1–53 (1999).

27. Jonasz, M. Size, shape, composition, and structure of microparticles from light scattering. in *Principles, Methods and Application of Particle Size Analysis* 143–162 (Cambridge University Press, 1991). doi:10.1017/CBO9780511626142.014.

28. Stetefeld, J., McKenna, S. A. & Patel, T. R. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys. Rev.* **8**, 409–427 (2016).

29. The principles of dynamic light scattering ∷ Anton Paar Wiki. https://wiki.anton-paar.com/at-de/das-prinzip-der-dynamischen-lichtstreuung/.

30. LS Instruments | Introduction. https://lsinstruments.ch/en/theory/dynamic-light-scattering-dls/introduction.

31. Ltd., M. I. *Achieving high sensitivity at different scattering angles with different optical configurations*. (2014).

32. Kim, A., Ng, W. B., Bernt, W. & Cho, N. J. Validation of Size Estimation of Nanoparticle Tracking Analysis on Polydisperse Macromolecule Assembly. *Sci. Rep.* **9**, 1–14 (2019).

33. Kim, A., Bernt, W. & Cho, N.-J. Improved Size Determination by Nanoparticle Tracking Analysis: Influence of Recognition Radius. *Anal. Chem.* (2019).

34. Picot, J., Guerin, C. L., Le Van Kim, C. & Boulanger, C. M. Flow cytometry: retrospective, fundamentals and recent instrumentation. *Cytotechnology* **64**, 109–130 (2012).

35. McKinnon, K. M. Flow Cytometry: An Overview. *Curr. Protoc. Immunol.* **120**, 5.1.1 (2018).

36. Flow Cytometry Guide - Creative Diagnostics. https://www.creative-diagnostics.com/flow-cytometry-guide.htm.

37. Yu, W. & Hancock, B. C. Evaluation of dynamic image analysis for characterizing pharmaceutical excipient particles. *Int. J. Pharm.* **361**, 150–157 (2008).

38. Whiting, J. G., Tondare, V. N., Scott, J. H. J., Phan, T. Q. & Donmez, M. A. Uncertainty of particle size measurements using dynamic image analysis. *CIRP Ann.* **68**, 531–534 (2019).

CHAPTER 3

A Compact Particle Size Analyser using Machine Learning

Author contribution: In this work I contributed to the design of the angular spatial filter, measurements for characterising and validating the system with the help of other researchers involved in the project. I did the initial data analysis and wrote the manuscript with my supervisor with contributions from all other authors. Further details of author contributions can be found in the above-mentioned article.

3.1 Introduction

Analysing the size and distribution of particles in a given sample plays a crucial role in improving and maintaining the quality of various products in many industries[1–3]. According to a report published by Research and Market, the global particle size analysis market is predicted to increase at a 5.8% compound annual growth rate (CAGR) from 371 million dollars in 2021 to 492 million dollars in 2026[4]. The growing demand for precise measurement instruments, particularly in nanotechnology applications, increased investment and expansion of R&D facilities in the pharmaceutical industry, and strict regulatory criteria for improved product quality across multiple industries can all be attributed to this growth.

The particle size analysis market can be divided into several categories based on the various measuring techniques used. Laser diffraction (LD) is the most used method, accounting for nearly 40% of the entire particle size analyser (PSA) market. PSAs based on LD are widely used because of their precision, high reproducibility, and fast measurement time, as well as their ability to measure a wide range of particle sizes from nanometres to millimetres. However, a number of shortcomings in current instruments are expected to hinder the expansion of the particle analysis market over the forecast period. As already mentioned in the previous chapter, these devices are expensive (~100–200 k€), heavy (~30 kg), and large (~690 × 300 × 450 mm) due to the fact that they incorporate several detectors (one sensor for each angle monitored). Moreover, a large distance between the samples and the detectors is necessary to obtain the required scattering

angular resolution. Another disadvantage of LD PSAs is the need to dilute samples above their typical concentration range of about 0.001–1% (v/v)[5], which requires additional sampling and dilution systems. This is due to the fact that these systems determine the particle size distribution (PSD) using Mie or Fraunhofer models[6,7], both of which are based strictly on single scattering events. However, in many industrial processes, it is necessary to measure suspensions at high concentrations, in which case multiple scattering[8,9] becomes a prominent issue. Although most commercial LD PSAs employ complicated multiple scattering correction algorithms[10–13] in their PSD analysis software, this increases the time-to-result and is not suitable for performing online measurements.

Machine learning (ML)[14] approaches are an alternate method for computing the PSD without the need of optical models or a sophisticated correction factor. ML is a powerful tool capable of adapting to process changes and delivering dependable outcomes through pattern recognition. It has previously been demonstrated in literature that using artificial intelligence, particle sizes can be determined for concentration ranges above the usual working range for commercial devices[15,16]. This approach can be particularly useful in many industrial applications, thus eliminating the need for time-consuming and tedious sample preparation steps. However, the high cost of these commercial devices limits the use of multiple sensors in the production line. This, in addition to the fact that they are high maintenance devices that require highly skilled staff to operate them, still remains a significant impediment to their use for online monitoring of industrial operations.

In this chapter, we introduce the concept of a novel, low cost, robust and miniaturised PSA using a collimated beam configuration, together with an image sensor and machine learning (ML), in particular, the random forest[17] (RF) algorithm. We will discuss the device's design and execution in depth, as well as how, depending on the particle size being measured, multiple scattering becomes noticeable at large concentrations. Lastly, we will demonstrate how the random forest algorithm can be used to correct this problem and achieve good size predictions with high accuracy.

3.2 Concept of PSA Using a Novel Angular Spatial Filter

Our newly developed PSA is built with off-the-shelf consumer electronic products, such as a complementary metal-oxide-semiconductor (CMOS) image sensor array and light emitting diode (LED) source. This significantly reduces the cost of the device (<10 k€) compared to the commercially available LD based devices. Moreover, unlike LD PSAs, which require several detectors placed at large distances in order to obtain an adequately resolved angular scattering profile, our device is extremely compact (on the order of 10 cm) as it only requires one CMOS image sensor array (camera), placed in close proximity to the sample. This is thanks to a key novel feature, namely a patented[18] very small (5 × 17 mm) micro-structured element made of plastic or resin, which we call the angular spatial filter (ASF). The ASF is made with an array of holes of different diameters and extruded from a polymer rod. As shown in **Figure 3.1 a** when particles are illuminated by a collimated light beam, each of these holes in the ASF acts as an aperture and allows collecting the forward scattered light by the particles up to a predetermined angle. This angle which we designate here as the cut-off angle, θ_c, can be calculated using geometrical optics (we neglect for simplicity diffraction effects) and is given by the following equation:

$$\theta_c = \arctan\left(D/L\right) \tag{3.1}$$

where, D and L are the diameter and the length of the filter hole, respectively. Note that θ_c is calculated for particles suspended in air. For particle suspensions in water, these angles need to

be corrected for the refraction encountered by the light rays in water. Hence, the actual scattering angle θ can be calculated using Snell's law:

$$\sin \theta = \frac{\sin \theta_c}{n_w} \tag{3.2}$$

where, θ_c is the detected scattering angle and n_w is the refractive index of water.

For the sake of simplicity, in this ASF design it is assumed that the response of each hole is like that of an ideal low pass filter, i.e., a rectangular response up to the corresponding θ_c. This means that the scattered light above θ_c will be completely absorbed by the side walls, there is no cross talk between the adjacent holes and there is no reflection from the inner walls of the ASF.
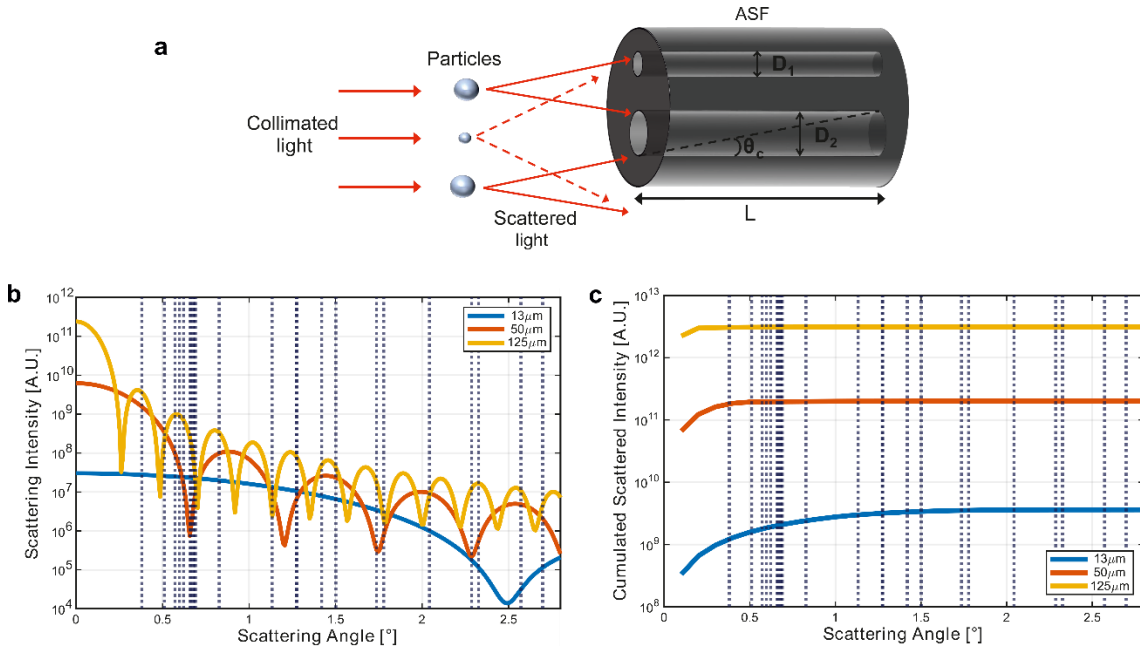
Figure 3.1: A novel ASF concept. (a) Schematic diagram of the ASF, showing dependence of the cut-off angle, θc, on the filter's length (L) and diameter (D) of the filter holes. (b) Simulation of the single particle scattering profile in water, for three different sizes of polystyrene particles using the Mie algorithm[19] in MATLAB. (c) Simulation of the cumulative scattering intensity for the three particle sizes in (b). The vertical dashed lines represent the cut-off angle for each ASF hole derived from equation (3.1) and corrected for refraction in water using (3.2).

In **Figure 3.1 b**, the scattering profiles of three sizes of polystyrene beads in water are simulated using the Mie scattering theory[19]. As mentioned previously, commercial LD PSAs rely on sampling such angular scattering patterns at each angle in order to determine particle size using a large number of sensors. In our PSA design, rather than measuring the scattered intensity at each angle, the ASF allows for the measurement of a cumulative scattering profile from zero to a defined cut-off angle (**Figure 3.1 c**). As a result, different groups of pixels can independently measure the light intensity from each hole of the ASF array when paired with a single CMOS camera. The cut-off angles that can be measured with our ASF are shown by the dashed lines in **Figure 3.1 b** and **c**. It is important to note here that for the calculation of the filter cut-off angles using equation **(3.1)**, we do not take into account the effect of the light diffraction caused by the circular apertures, i.e., the filter holes. We assume that the ASF is still able to perform angle resolved light-power measurements and that the PSA can distinguish particle size and concentration.

3.3 Fabrication of the ASF

The aspect ratio of the holes, which is described by the ratio of their length to their diameter, L/D, is a major concern in the fabrication of the ASF. In order to measure particle sizes from tens to hundreds of microns, it is necessary to sample small scattering angles, as large particles scatter light mostly in the forward direction. For example, if we look at the scattering profile of the 125 µm particle in **Figure 3.1 b**, the principal lobe has a minimum at around 0.2°. Using (**3.1**), the L/D ratio can be calculated to be 286. Such a high aspect ratio feature is extremely challenging to fabricate as it means, for example, fabricating a hole of 60 µm over a length of 17 mm. Even the most advanced 3D micro-printers on the market are unable of layer-by-layer growth of such small features over many millimetres. Maskless photolithography, for example, can achieve submicron resolution but cannot produce structures with such a high aspect ratio. Other approaches combine additive manufacturing with micro machining. However, these increase the final cost of the piece and have never achieved optimal results. For example, laser sintering, selective laser melting and laser drilling can improve small feature resolution in the order of microns but this would mean using multiple parts that require proper alignment.

A different approach that can be used for fabricating the ASF filter is polymer extrusion, which is often used for producing micro-structured polymer optical fibres (mPOFs)[20]. This technique has allows alleviating the above-mentioned fabrication problem posed by conventional methods, leading to a highly customised large number of holes with high aspect ratio. To make the ASF, a preform of 60 mm in diameter and 100 mm in length, made from commercially available poly(methyl methacrylate) (PMMA) rods is prepared by drilling holes with a computer numerical control (CNC) mill, using specially coated drill bits with a size range from 1 mm to 10 mm in diameter. This primary preform is then annealed at a temperature of 80 °C for a week and then drawn to obtain a micro-structured secondary preform or 'cane' with a reduced diameter and a length of a few centimetres. Details of the entire fabrication method can be found in reference [21]. This fabrication process allows for a lot of flexibility in designing the ASF as the diameter and length of the filter may be easily modified to collect the scattering angles required for certain applications.

The first prototype of the ASF is designed to measure particle sizes in the range from 10 µm to 130 µm, approximately. Hence, the micro-structured preform is drawn to canes of length 50 mm and diameter 5 mm to fit well within the CMOS image sensor area. The ASF is 17 mm long, consisting of 23 holes with diameters ranging from 112 µm to 800 µm. This combination of diameter and length allows the collection of scattering angles from 0.38° to 2.7°. Using equation (**3.2**), the actual angles are found to be from 0.29° to 2.02°, and using Mie theory, we confirmed that this angular range is suitable for measuring our preferred particle size range, i.e., 10-130 µm.

Since the ASF is manufactured from PMMA, it is mostly transparent in the visible wavelength range. For the concept of the ASF to work, it is necessary that the side walls of the filter are absorbing. To achieve this, the filter is coated with black acrylic ink, which also helps to reduce reflections from the walls, and minimises crosstalk between neighbouring holes. The outside of the filter is also covered with a black tape.

As mentioned previously, multiple scattering becomes a dominant effect when measuring particle suspensions at high concentrations. To measure this, we need to collect angles larger than 2.02° with our ASF, as the multiple scattering effect causes an apparent increase in the scattering angle. A clever way to collect large scattering angles is by polishing one side of the ASF along the whole length to give it a semi-circular shape and then inserting it in a holder the same length as the filter (**Figure 3.2**). The remaining semi-circular space in the holder acts as a big aperture, in theory collecting all of the sample's forward scattering angles.
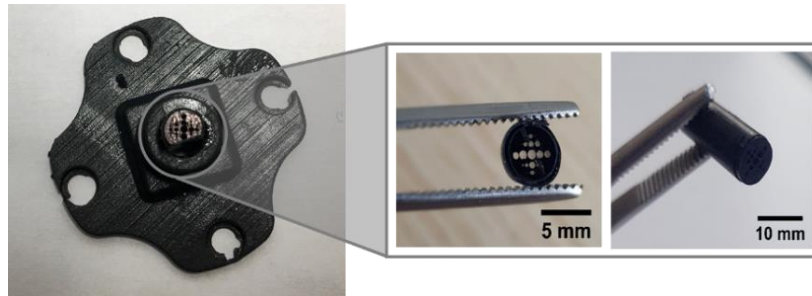
Figure 3.2. Angular spatial filter (ASF). *It is fabricated from PMMA and coated with black acrylic ink. The ASF is 17 mm long and consists of 23 holes of various diameters.*

3.4 Proposed PSA Design

To construct the PSA based on the ASF, we used a fibre-coupled LED at a wavelength of 632.8 nm, together with a collimator to illuminate the samples with a collimated beam of 10 mm in diameter. The ASF, together with the holder, is attached to the CMOS camera and placed in close proximity to the sample in order to collect both the forward unscattered and the scattered light. The CMOS camera (Micron MT9P0311) has an active area of 24 mm^2 (5.7 mm in length and 4.28 mm in width). As the ASF is placed close to the flow cell, the active area of the CMOS is also the field-of-view of the system. The CMOS camera has 2592 × 1944 pixels, with each pixel being 2.2 × 2.2 µm in size, and a frame rate of 14 frames per second is used to obtain a full-resolution image. The sensor consists of four colour channels but only the red channel is used for image processing. A schematic diagram of the proposed PSA design, together with the lab prototype, is shown in **Figure 3.3**.
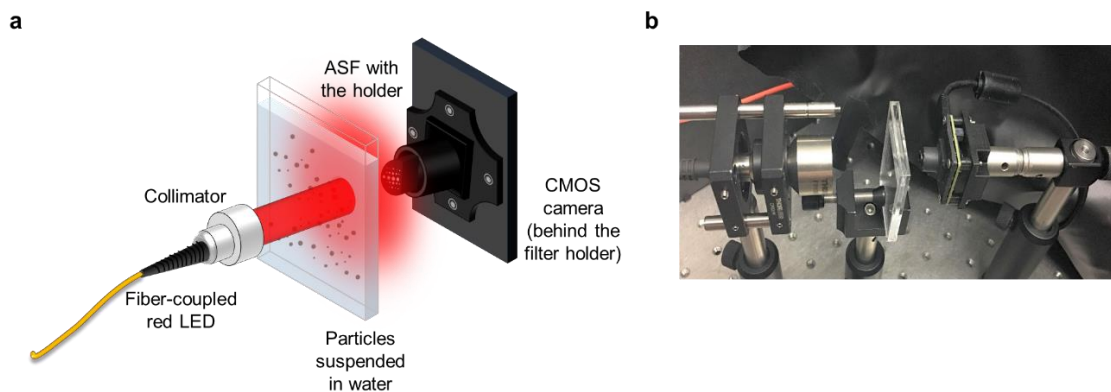
Figure 3.3*: **Design of the proposed PSA. (a)** The PSA is shown schematically with the new ASF coupled to the CMOS camera and a collimated LED source. **(b)** Lab-scale prototype of the PSA. In both (a) and (b) particle suspension is shown to be in a cuvette. In order to maintain homogeneity, magnetic stirrers have to be used, which are not shown in the figures.*

3.5 Measurement of Particle Suspensions Using the PSA

3.5.1 Experimental Set-Up

The samples tested with the proposed PSA were glass beads of various size distributions, dispersed in water to prepare different concentrations. The beads have a refractive index of 1.51

at the working wavelength of the PSA. The particle characteristics with the concentrations tested are listed in **Table 3.1**.

Table 3.1: Characteristics of the measured samples.

Sample	Size range (µm)	Density (g cm⁻³)	Commercial LD PSA			Concentrations tested (mg mL⁻¹)
			D10 (µm)	D50 (µm)	D90 (µm)	
Guyson	80	2.5	55	74	92	1,5,10,15,20,25,30,40,50
	40		24	39	56	1,5,10,15,16,18,20,25,30
Cp500	13-20	2.56	6	11.9	21	1,2,3,4,5,6,7,8,9,10
Sovitec	0-50	2.46	18	34.8	51	1,5,10,15,18,20,22,25,30,40
	40-50		33	43.6	51	
	40-70		46	62.3	80	
	70-110		68	87.5	108	
	90-150		97	125.5	157	

The glass beads were initially characterised using a scanning electron microscope (SEM). To prepare the particle solution, a known mass of the glass beads in powder form was dispersed in a known volume of deionized (DI) water in order to prepare the initial stock solution with a concentration of 100 mg mL⁻¹, from which the different concentrations were prepared. The stock was stirred continuously at 300 rpm with an overhead stirrer to maintain a homogeneous suspension in the beaker.

A schematic diagram of the entire experimental setup is illustrated in Error! Reference source not found.. Dark images were acquired first without the LED light and, subsequently, subtracted from the images acquired with the samples. Prior to measuring the samples, 200 mL of DI water was circulated through the flow cell using the peristaltic pump, and five images, which we call the reference images, were captured with a time delay of 20 to 60 seconds between each image. It is important to control the pressure of the peristaltic pump in such a way that no bubbles form while the sample passes through the flow cell. Next, a known volume of the sample stock was added to the water and images were captured. The same samples were also measured with a commercial laser diffraction device from Sympatec in order to determine the size distribution. Before measuring each set of samples, the flow cell was thoroughly cleaned by flowing DI water through it several times. All the images are analysed using the 'regionprops' function in MATLAB to, firstly, locate the holes within the images and then calculate their corresponding intensities.
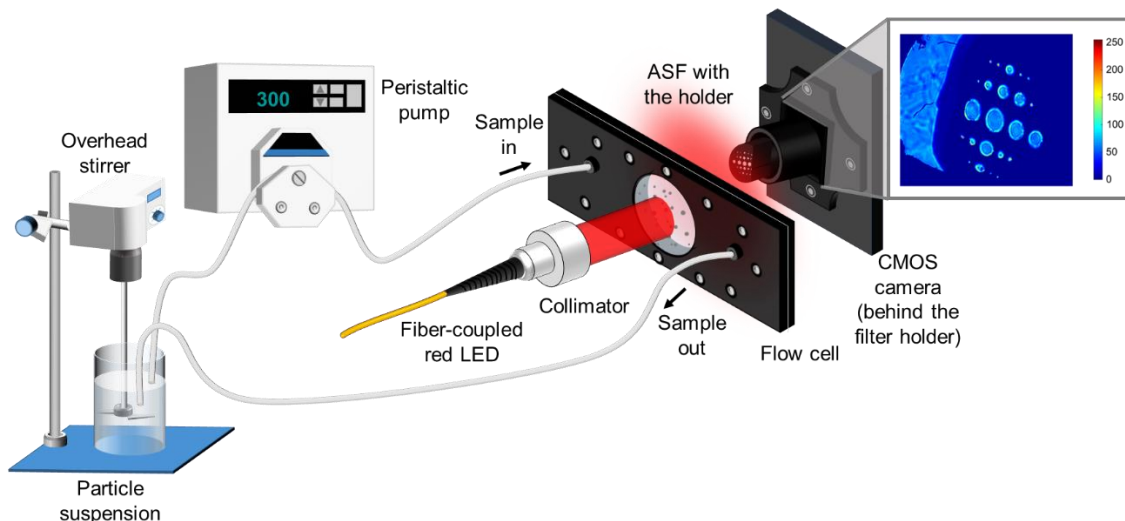
Figure 3.4: Experimental set-up. The schematic diagram depicts the flow cell for circulating the sample using a peristaltic pump. An overhead stirrer is used to prevent agglomeration and deposition of the particles at the bottom of the beaker. A raw image from the CMOS camera for Sovitec 90-50 µm particle at 10 mgmL^{-1} is also shown.

From the raw image, it can be observed that in some holes the images show local intensities deviating from the corresponding average values. These are likely to be associated with residual diffraction and reflection from the inner part of the hole walls, as well as geometrical and material imperfections of the ASF (e.g., imperfect hole geometry, slight misalignment between holes and missing black ink). Next, we will show that despite these imperfections and associated effects, the ASF angular dependence of the transmission of each hole is maintained and the ASF is still very efficient in discriminating different particle sizes and concentrations.

3.5.2 Effect of Multiple Scattering on Particle Measurement

For a fixed particle size, the light distribution in the ASF holes depends on particle concentration. This is shown in **Figure 3.5 a**, where we plot the average intensities in the filter holes, normalised to those of the reference (water) against the filter cut-off angles at three different concentrations for the particle size range of 40-50 µm. For a fixed concentration, **Figure 3.5 b** demonstrates that smaller particles have a significantly stronger effect, not just on the scattered intensity, but also on the angular dependence. This phenomenon can be explained in terms of the multiple scattering effect, which causes the particle's scattering lobe to widen, resulting in a decrease in the scattered intensity detected by the CMOS camera. This is also confirmed by the graph in **Figure 3.5 c**, where the average scattered intensity in the smallest ASF hole against concentration is plotted for three different particle size distributions. For 13-20 µm particles there is a sharp decrease in the intensity after a certain concentration value. This decrease is because light is scattered off from additional particles at larger scattering angles that exceed the cut-off angle of the filter holes. Since large particles scatter light at very small angles, the multiple scattering effect is less prominent for these even at fairly high concentrations. This is evident from the curve for 90-125 µm particles.
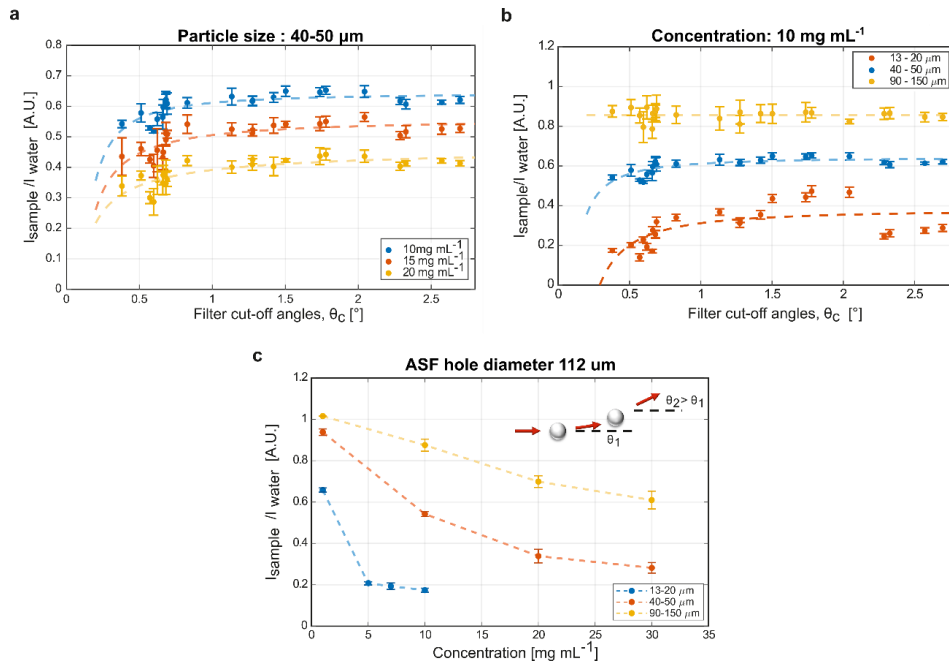
Figure 3.5: Multiple scattering effect (a) *The average intensities of the filter holes, normalised to those of water, for glass beads with size distribution 40–50 µm are plotted as a function of θc for three different concentrations. **(b)** The average intensities of the small filter holes are plotted as a function of the filter cut-off angles (θc) for the three ranges of glass bead diameters with the same concentration (10 mgml⁻¹). In **(a)** and **(b)** the error bars represent 95% confidence interval and the dashed lines represent a least square fit. **(c)** The average intensity, normalised to water, of the 112 µm diameter hole against concentration for three different glass bead diameter distributions, 13–20, 40–50 and 90–150 µm. The dependence on concentration, which increases with the decreasing size of the glass beads, is a sign of multiple scattering.*

3.6 Machine Learning Algorithm for Data Analysis

Most commercial LD PSAs define a working concentration range depending on the particle sizes to be measured. This is because too low a concentration results in a low scattering intensity being measured, leading to a poor signal-to-noise ratio and unreliable results, while too high concentration gives rise to multiple scattering, meaning that the particle size is underestimated. Hence, in order to overcome this problem of concentration dependence and to be able to measure particle sizes with a wide concentration range, we developed a machine learning algorithm based on the RF algorithm.

The RF algorithm is a supervised learning algorithm in the sense that it is trained to make predictions by observing patterns in a given dataset. This model can be sub-divided into classification or regression algorithms. The RF regression approach is used in this study to predict continuous values such as particle sizes. To explain briefly, the "forest" in RF model refers to an ensemble of decision trees, each of which makes its own prediction and the final outcome of the model is determined as the average of the predictions from all the decision trees. This averaging improves model precision while reducing overfitting[22]. A root node, decision nodes, and leaf nodes make up each decision tree in an RF algorithm. The root node represents the entire dataset and is the top node in a tree. The dataset is sent down from the root node to the leaf nodes via decision nodes during the training process through recurrent splitting based on particular attributes and their threshold values until the leaf nodes are all homogeneous and can no longer be split further. More details on RF algorithm can be found in references[17,22]. **Figure 3.6** below depicts the RF structure along with an example of a simple decision tree showing the different nodes.
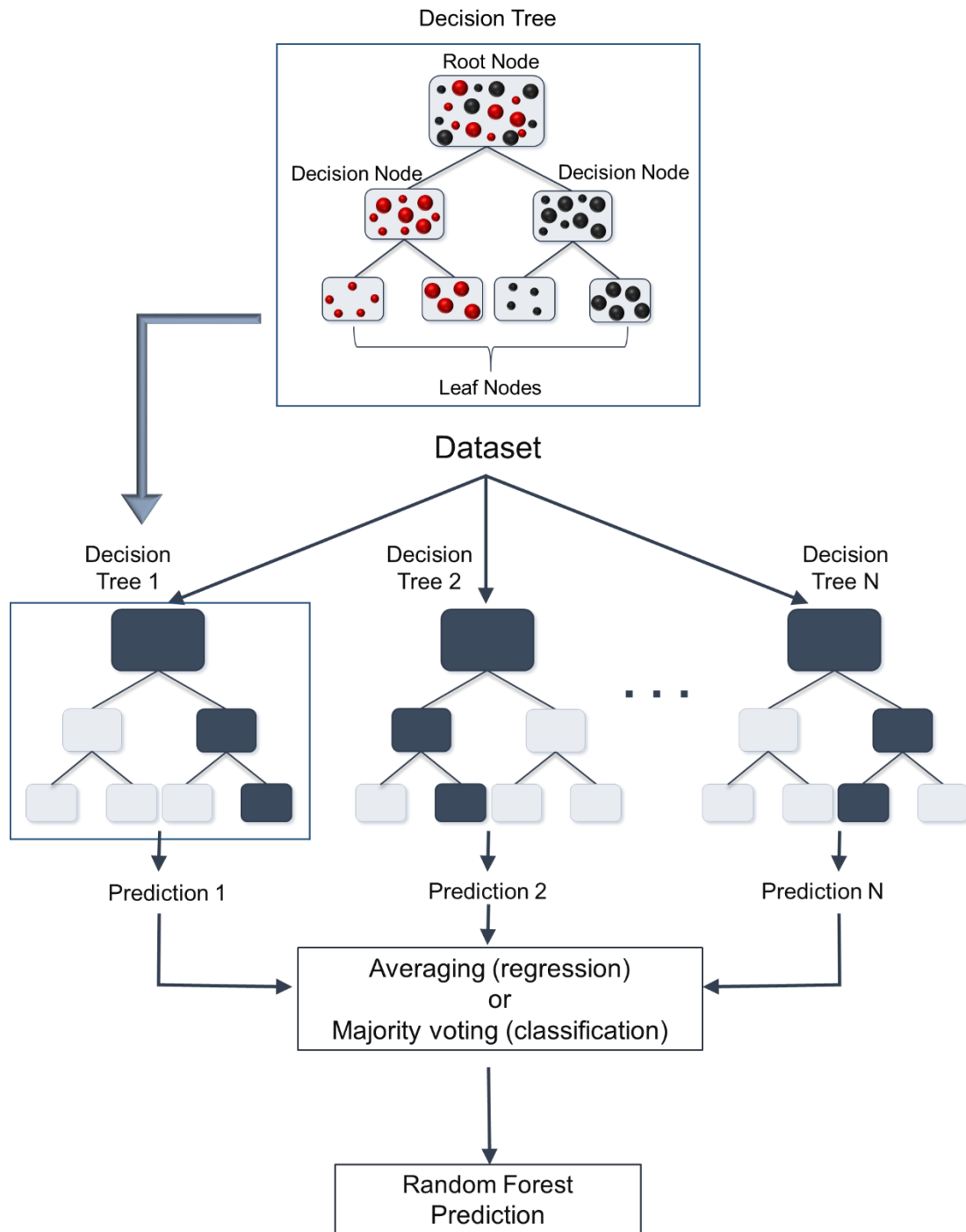
Figure 3.6: Random Forest (RF) Algorithm. *A simple example of a decision tree showing the root node, decision nodes and the leaf nodes (**top**). Structure of the RF algorithm (**bottom**). Multiple decision trees make up the RF model. Each tree represents a decision-making model in the form of a tree. Each decision (data split) is based on a single feature and its threshold value. The features, threshold values, and when to end the tree are all part of the training process.*
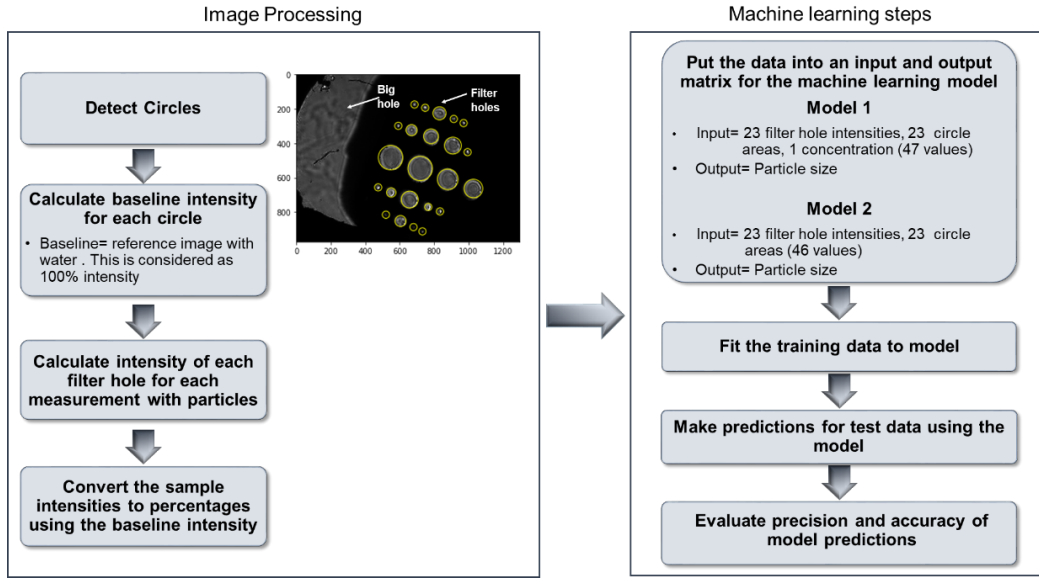
Figure 3.7: Flowchart showing the steps of the machine learning algorithm used for prediction.

The steps of the algorithm used for predicting particle are shown in **Figure 3.7** and explained in detail below:

1. The first step is similar to the analysis carried out in the previous section. The location of 23 filter holes in the sample images are determined using an image processing library (scikit-image, blob detection) in Python, and then the pixel intensities inside the filter holes are calculated, together with their corresponding diameters. These intensity values are then normalised with respect to the reference images and are expressed as relative intensities in percentages.

2. The entire data set, consisting of 459 images, is randomly divided into two sets, namely the training set (344 images) and the testing set (115 images). Each image is from a particle suspension with known volume median diameter D50 value and concentration. Relative intensities, hole diameters and concentrations are given as input to the ML model and particle sizes are set as the target values.

3. In order to find the correlation between the input and the output parameters, we tested two algorithms suitable for structured data like ours from among the various ML algorithms available[23], these being the gradient boosting and the RF algorithms. This preliminary test showed that the RF algorithm gives a slightly better prediction for the number of data points used in the analysis. A scikit learn machine learning library was used to create the model.

4. Since the training and testing sets are randomly divided, different splitting combinations of these sets can cause variations in the model performance. To account for this fluctuation, we first observed the mean and the standard deviation of the model predictions from one train-test split and evaluated the model performance using the mean absolute percentage error (MAPE) for the model predictions. The MAPE for *n* number of images is given by the following equation:

$$MAPE = \frac{100\%}{n} \sum_{i=0}^{n} \left| \frac{Actual\ value_i - Predicted\ value_i}{Actual\ value_i} \right| \quad \text{(3.3)}$$

Then the model is tested again on a different randomly-split combination of training and testing sets. After repeating this process 100 times no significant change is observed in the model performance. Therefore, the model is trained and tested 100 times and the average MAPE of all these tests, together with their standard deviations, are reported as the figure of merit.

The algorithm tested up to this point is referred to as Model 1. For real applications it is necessary for the PSA to be able to determine the particle size, i.e., D50, without any input concentration. Hence, we developed another model, Model 2, to predict the D50, where the input parameters used are only the relative intensities in the filter holes and their respective diameters.

3.7 Particle Size Prediction Using the Developed ML Model

The predictions obtained from a single test set for Model 1 are shown in **Figure 3.8 a** and **b**. The average MAPE from 100 tests is found to be 2.52%, with a standard deviation of 0.73%. It can be seen from **Figure 3.8 a** that there is good agreement between the predicted and the nominal particle size using the RF model. It is also evident that the model can correct the effect of multiple scattering caused by concentration dependence, as shown in **Figure 3.8 b,** where the particle size predictions remain constant with changing concentration.
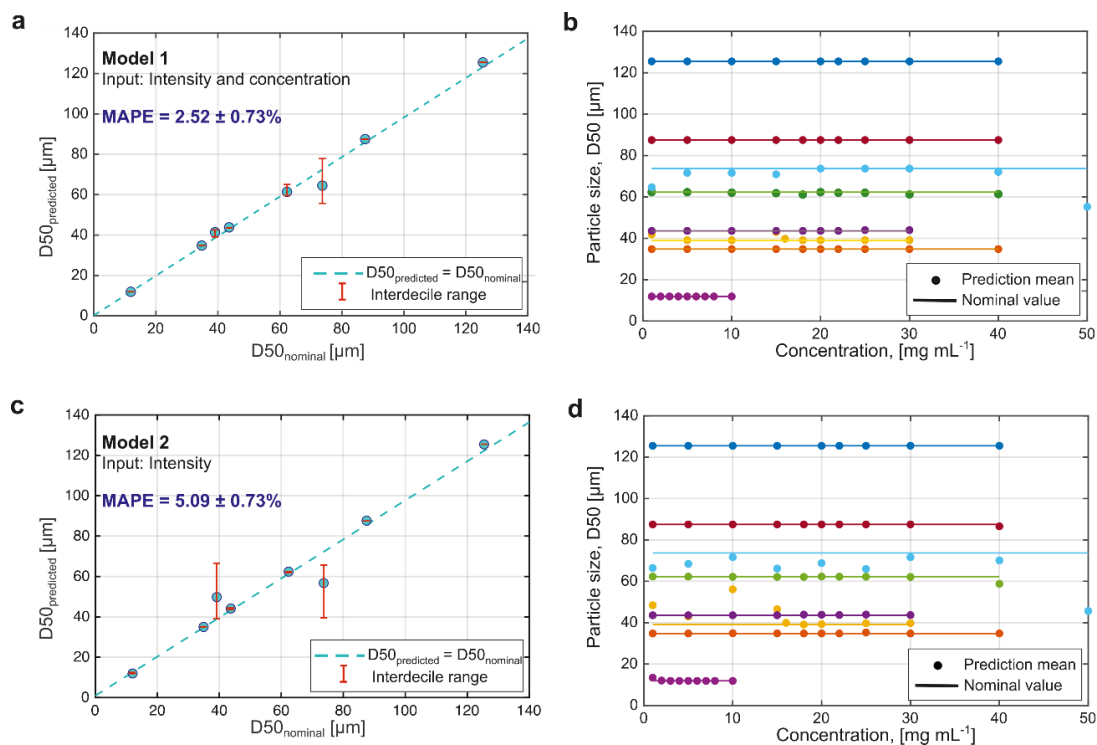
*Figure 3.8: Particle size prediction using Model 1 and Model 2. The mean predicted D50 values for one of the test sets plotted against (**a**) the nominal D50 values and (**b**) particle concentration using Model 1. Despite the multiple scattering effects, the predicted diameters are close to the nominal diameters (straight lines). (**c**) The mean predicted D50 against nominal diameter and (**d**) the D50 prediction against concentration using Model 2. The dashed line represents predicted diameter = nominal diameter. The interdecile range is also shown for each predicted D50.*

When tested with Model 2 (**Figure 3.8 c** and **d**), the mean prediction error of 5.09% is found to be higher than that with Model 1. This is as expected since, in this model, concentration was not used as an input parameter to train the model.

However, we observe a larger deviation from the nominal size for particles with diameters 39 and 74 µm, as can be seen in **Figure 3.8 d**. A closer look at the microscope images (**Figure 3.9 a** and **b**) shows that there are some non-spherical particles present in these samples. The random shape and orientation of these particles may have an effect on the angular distribution of the scattering intensity. Therefore, we tested Model 2 again without including the measurements from these samples (**Figure 3.9 c**) and then once more with only these two samples (**Figure 3.9 d**).
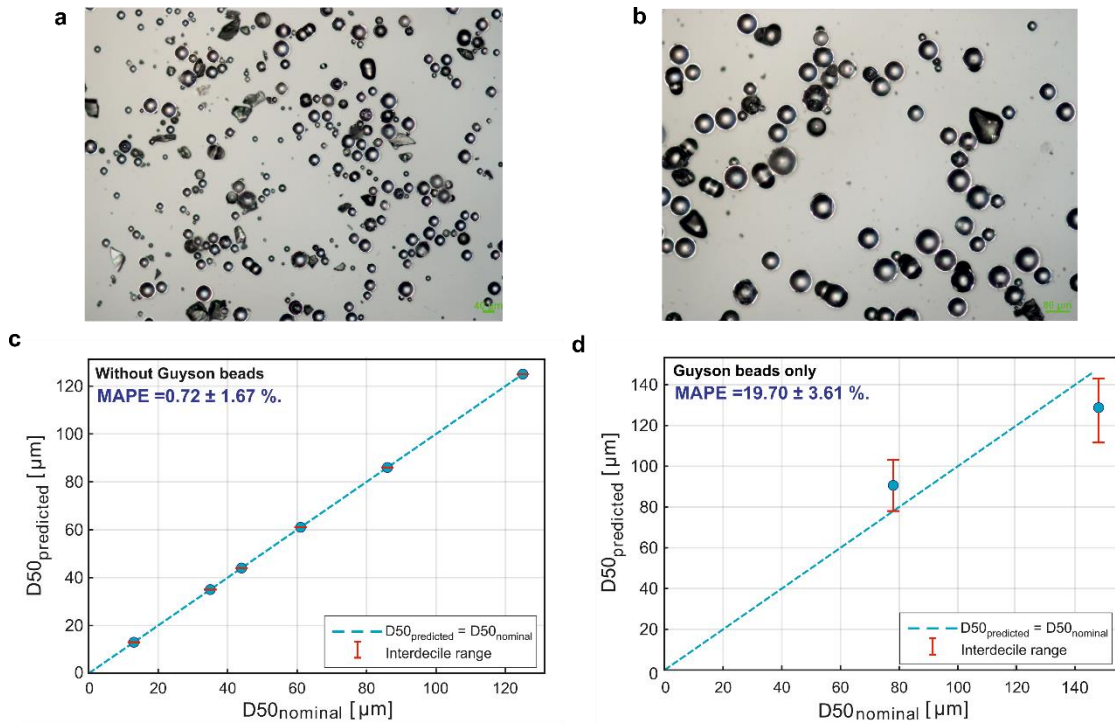
Figure 3.9: Particle size prediction with and without non-spherical particles. Microscope image of Guyson beads with a D50 of (a) 39 µm and (b) 74 µm show the presence of some non-spherical particles. Performance of Model 2 (c) without Guyson beads and (d) with Guyson beads only. It can be seen from (c) that when the model is trained and tested without Guyson beads, the model performance improves significantly whereas in (d), with only Guyson beads, the MAPE is seen to increase.

The significant improvement in the prediction (0.72%) compared with that obtained without these samples confirms that it was indeed the presence of these particles with non-spherical shape that increased the MAPE for Model 2, particularly in the first test.

In some practical applications, it is necessary to determine the D10 and D90 percentile values, in order to give an idea about the size of the fine and coarse particles present in a sample. Hence, the next test we performed involved predicting the size distribution of the measured samples. This process can be performed very easily with our proposed PSA, without any system adaptation or modification, by simply training the ML model with not only the D50 value but also the D10 and D90 percentile values. If necessary, additional percentiles, such as D5, D15 or D95, can also be included in the machine learning training. On testing Model 1 with D10, D50 and D90 values, the MAPE was found to be $4.27 \pm 1.64\%$, $3.02 \pm 1.07\%$ and $2.4 \pm 0.8\%$, respectively. Note that these values were obtained with only one set of D10 and D90 data for each size, but the prediction accuracy could be further improved by training the model with more data from samples with the same median diameter but a varying size distribution.

In order to further assess the functionality of the ASF and show that the information obtained from the different sizes of ASF holes is crucial for the device and the ML model to work, we performed an additional test using only the big hole intensity from the same images used for the ASF holes. As mentioned previously, the big hole allows the entire angular distribution of the forward scattered light to be collected from the sample, thus providing the absorption information of the particle solution too. For the ML analysis, Model 1 requires 47 input values, so to train and test the model we used intensity values from the big hole within regions defined by the 23 filter hole areas (**Figure 3.10 a**).
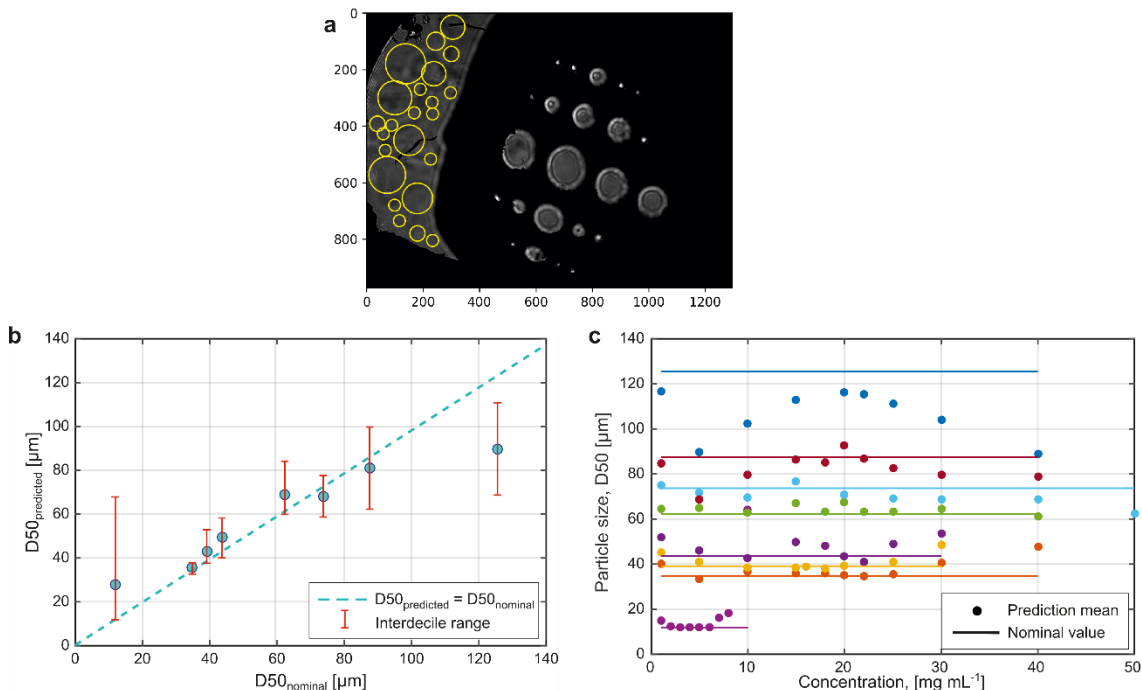
Figure 3.10: Testing Model 1 with big hole intensity and concentration as input. The yellow circles in (a) represent the 23 filter holes placed in a random manner inside the big hole region. The mean predictions against nominal values for one of the test sets are shown in (b) and against concentration in (c).

The MAPE of 23.03 % (**Figure 3.10 b**) and the deviations from predictions with respect to nominal values for varying concentrations are significantly larger than those obtained with the ASF filter (**Figure 3.8 a** and **b**). This clearly indicates that scattering selectively measured through ASF filter is required for the machine learning model to predict particle sizes with a high degree of accuracy. Also, this analysis, together with the flow-through results reported in the next section, confirms the absence of any correlation between the images used to train the ML model other than those solely related to particle size and concentration.

3.8 Flow-Through Measurements

In many industries, monitoring system changes rather than exact values is necessary for online operations. In order to demonstrate the capability of the device and the machine learning model for flow-through measurements, rather than batch measurements, we performed preliminary tests with two samples, 13-20 μm (Sample 1) and 40-70 μm (Sample 2). In the first trial series, Sample 1 was measured at a concentration of 2 mgml-1, followed by Sample 2 at 5 mgml-1, without cleaning the flow cell. The same experiment was then repeated later the same day. For the second trial series, Sample 1 was measured at three different concentrations, one after the other, followed by Sample 2, again at three different concentrations, one after the other. The same experiment

was then repeated, but this time, Sample 2 was flowed first, followed by Sample 1. The previous machine learning Model 1 was then calibrated with three of these new data sets and tested on a fourth data set. The MAPE for Model 1, using only two samples, was found to be $1.77 \pm 0.25\%$, as shown in **Figure 3.11**, showing the potential of the current PSA design and the ML model in predicting particle size changes in real-time. With further system modification and optimisation, the prediction error could be lowered, further improving the precision for such flow through measurements.
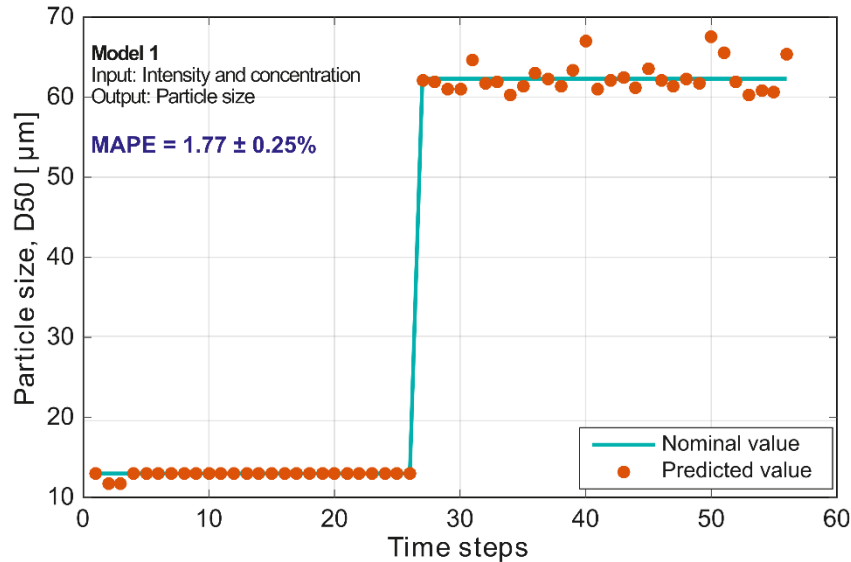
Figure 3.11: Flow-through measurement. Two samples, 13-20 μm (D50 13 μm) and 40-70 μm (D50 62.3 μm), are measured continuously, one after the other, without cleaning the flow cell. Note that with the current set-up, the transition from one particle size to another is too fast to capture a sufficiently large number of images to properly train the ML model. For this reason, the graph does not contain a transitory region where both particle sizes coexist.

3.9 Conclusion

In this chapter, we have presented a portable, low-cost and, most importantly, compact PSA, using mostly off-the-shelf components. The real innovation in the proposed PSA is a patented small hole structure called angular spatial filter (ASF), which enables characterisation of the scattering properties of wet dispersions without the need for multiple detectors of commercially available systems. The ASF, when combined with an LED and a CMOS camera, allows us to capture images that we then use with our custom-designed machine learning algorithm to predict particle's median diameter. To validate the proposed PSA, we measured glass beads of various size distributions at several concentrations. We have shown how multiple scattering affects the angular scattering pattern and scattering intensity of different particle sizes, and how our developed ML model can correct this concentration dependence. The prediction error, MAPE, obtained with concentration as an input parameter was 2.52% while without concentration the error increased to around 5%. This was due to the presence of non-spherical particles in two samples, which, when removed from the analysis, reduced the MAPE to only 0.72%, without predefining concentration as an input. This value, determined using polydisperse particles, is well within that recommended by international ISO standard 13320[24] for LD PSAs, and is also comparable to the accuracy of 0.6% for monomodal latex standards specified by the commercially available laser diffraction-based devices. We have further demonstrated the potential of the PSA for online or atline applications by performing a proof-of-concept flow-through measurement with two different particle sizes and concentrations. Although the error limit of such measurements using

the PSA is not as low as that achieved by commercial LD based devices, in online operations, for example milling, such devices are often used to monitor system changes rather than measuring exact values. As a result, the performance is a trade-off between the device's low cost and the desired accuracy. Additionally, as online analysis reduces sample intrusion, measurements in many particle processes are often more representative of the system, despite larger absolute device error.

The portability and low cost, combined with a powerful machine-learning model, makes our proposed PSA an attractive solution for different industrial processes without the need to perform complex sampling and dilution operations. This opens up opportunities for our novel PSA in a wide range of applications. In the next chapter we will demonstrate that since the PSA is sensitive to changes in the refractive index between the measured particle and the surrounding medium, it can also be used for relevant biological applications, such as the detection of microorganisms (e.g., *Escherichia coli*), in water.

3.10 References

1. Valsangkar, A. J. Principles, methods and applications of particle size analysis. *Can. Geotech. J.* **29**, 1006–1006 (2008).

2. Shekunov, B. Y., Chattopadhyay, P., Tong, H. H. Y. & Chow, A. H. L. Particle size analysis in pharmaceutics: Principles, methods and applications. *Pharm. Res.* **24**, 203–227 (2007).

3. Servais, C., Jones, R. & Roberts, I. The influence of particle size distribution on the processing of food. *J. Food Eng.* **51**, 201–208 (2002).

4. Global Particle Size Analysis Market by Technology (Laser Diffraction, DLS, Imaging, Coulter Principle, Sieving, Nanoparticle Tracking), Dispersion (Wet, Dry, Spray), End-user (Pharma-biotech, Cosmeceutical, Chemicals, Food, Academia), and Region - Foreca. https://www.researchandmarkets.com/reports/5116160/global-particle-size-analysis-market-by?utm_source=GNOM&utm_medium=PressRelease&utm_code=qjrsvx&utm_campaign=1629282+-+Global+Particle+Size+Analysis+Market+(2021+to+2026)+-+Increasing+Number+of+Conference.

5. Merkus, H. G. Particle Size, Size Distributions and Shape. in *Particle Size Measurements* 13–42 (Springer Netherlands, 2009). doi:10.1007/978-1-4020-9016-5_2.

6. Vargas-ubera, J., Aguilar, J. F. & Gale, D. M. Light-Scattering Patterns Using Three Inversion Methods. (2007).

7. Ye, Z., Jiang, X. & Wang, Z. Measurements of particle size distribution based on Mie scattering theory and Markov chain inversion algorithm. *J. Softw.* **7**, 2309–2316 (2012).

8. Mishchenko, D. Travis, L. & Lacis, A. *Multiple Scattering of Light by Particles: Radiative Transfer and Coherent Backscattering*. (Cambridge University Press, 2006).

9. Lenke, R. & Maret, G. *Multiple scattering of light: Coherent backscattering and radiative transfer*. (Cambridge University Press, 2000).

10. Gomi, H. Multiple scattering correction in the measurement of particle size and number density by the diffraction method. *Appl. Opt.* **25**, 3552 (2009).

11.   Quirantes, A., Arroyo, F. & Quirantes-Ros, J. Multiple light scattering by spherical particle systems and its dependence on concentration: A T-matrix study. *J. Colloid Interface Sci.* **240**, 78–82 (2001).

12.   Wei, Y., Shen, J. & Yu, H. Numerical calculation of multiple scattering with the layer model. *Particuology* **7**, 76–82 (2009).

13.   Harvill, T. L. & Holve, D. J. Method for measuring particle size in the presence of multiple scattering. (1997).

14.   LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).

15.   Guardani, R., Nascimento, C. A. O. & Onimaru, R. S. Use of neural networks in the analysis of particle size distribution by laser diffraction: Tests with different particle systems. *Powder Technol.* **126**, 42–50 (2002).

16.   C.A.O. Nascimento a. Use of neural networks in the analysis of particle size distributions by laser diffraction. *Powder Technol.* **90**, 89–94 (1997).

17.   Breiman, L. Random Forests. *Mach. Learn. 2001 451* **45**, 5–32 (2001).

18.   Pruneri, V., Martïnez Cordero, P. A. & Jofre Cruanyes, M. Apparatus for measuring light scattering. (2018).

19.   Bohren, C. F. & Huffman, D. R. *Absorption and scattering of light by small particles*. (John Wiley & Sons, 2008).

20.   Barton, G., Van Eijkelenborg, M. A., Henry, G., Large, M. C. J. & Zagari, J. Fabrication of microstructured polymer optical fibres. *Opt. Fiber Technol.* (2004) doi:10.1016/j.yofte.2004.05.003.

21.   van Eijkelenborg, M. *et al.* Microstructured polymer optical fibre. *Opt. Express* **9**, 319 (2001).

22.   Hastie, T., Tibshirani, R. & Friedman, J. Random Forests. in 587–604 (Springer, New York, NY, 2009). doi:10.1007/978-0-387-84858-7_15.

23.   Hastie, T., Tibshirani, R., Friedman, J. & Franklin, J. The elements of statistical learning: data mining, inference and prediction. *Math. Intell.* **27**, 83–85 (2005).

24.   13320. Particle size analysis-laser diffraction methods. *ISO Stand. Auth.* (2009).

CHAPTER 4

PSA for Biological Applications

Author contribution: My contributions to this work include modifying the PSA, preparing the particle solutions, and measuring them. I carried out the initial E. coli and Enterococcus sp. experiments with the assistance of other researchers in the group. I analyzed the raw data and performed the machine learning analysis, using codes written by a collaborator on this project.

4.1 Introduction

Rapid detection of pathogenic bacteria is essential in preventing outbreaks of many serious diseases[1,2]. The traditional method employed widely for identifying and monitoring bacteria relies on culturing and counting colonies[3] formed by microorganisms on agar plates. This method is time-consuming and involves tedious sample preparation steps. The incubation time required for visible colony formation by most foodborne or waterborne pathogens is usually 24 to 36 hours and can even be 72 hours, depending on the target strain. In many practical applications, such as detecting microbial contamination in the food processing industry[4,5], or monitoring wastewater treatment plants[6,7] and recreational beaches[8], timely detection is of paramount importance as delayed results can lead to public health hazards. Other less standard detection techniques like molecular-based polymerase chain reaction (PCR)[9,10], and immunological techniques such as enzyme-linked immunosorbent assay (ELISA)[9,11], etc., may provide faster analysis compared to conventional culture-base methods, but they are expensive and require sample pre-treatment, skilled personnel and a specialised laboratory.

Optical biosensors based on Mie light-scattering[12–14], on the other hand, have gained much popularity as an interesting alternative for pathogen detection, due to their short time-to-result, low cost, ease of use and non-destructive nature. Moreover, they allow true label-free and reagentless detection, in the sense that they do not rely on binding between target biomolecules and some recognition elements, like enzymes, antibodies etc., needed in other label-free sensors[15–17]. In light-scatter based devices, the angular scattering signature obtained from biological microparticles can be used to infer their size, shape and refractive index in order to identify their type. Simultaneously, for a specific type of microorganism, their concentration in a sample can be determined by measuring the change in intensity scattered by the cells. These sensors are promising detection tools as they allow real-time, onsite measurement with minimal sample preparation steps.

In recent years, these light-scatter based technologies have combined the benefits of machine learning (ML) to develop portable and cost-effective platforms for the detection and identification of microorganisms[17]. In such systems, the scattering pattern obtained either from bacterial colonies grown on agar plates[18,19] or in liquid suspensions[12,20] are used to classify different types

of bacteria. We have used a similar approach to investigate the application capability of our ML-based PSA in detecting and classifying microorganisms, particularly *Escherichia coli* (*E. coli*) and *Enterococcus sp.* in suspension. These bacteria were chosen because they are among the most common indicators of faecal contamination[21] which can result in serious illnesses. Apart from that, *Enterococcus sp.* is a pathogen that can develop biofilms on surfaces of medical devices or catheters, and it is one of the primary causes of various infections such as urinary tract infection, bacteremia, endocarditis, prostatitis as well as wound infections[22]. *E. coli* are Gram-negative rod-shaped bacteria with an approximate width of 0.5 µm and can vary in length from 1.0 to 3.0 µm[23]. In contrast, Enterococcus *sp.* are Gram-positive bacteria having spherical or oval cells with diameters ranging from 0.5 to 1 µm[24]. Most often, these bacteria occur in pairs or short chains. Because bacteria have somewhat higher refractive indices than water, 1.38 to 1.4 Mie-scattering patterns can be used to detect their presence directly.

The scattering signal measured by our PSA is dependent on sample concentration. Hence, we can use this property to determine the concentration of *E. coli* and *Enterococcus sp.* in water with the assistance of ML. This chapter introduces the design of a modified PSA for measuring small particles the size of bacterial cells and shows how the system is made more sensitive to small particle measurements compared to the previous setup. We then demonstrate how, by using ML algorithms, our PSA can classify between two different types of bacteria (*E. coli* and *Enterococcus sp.*) and particles and can also predict their concentration in water.

4.2 Optimization of the PSA for Small Particle Measurement

In this study, some modifications have been made to the ASF to reduce residual reflections from the inner walls of the ASF observed in previous measurements. Although the angular dependence was maintained in the previous filter, to further improve the functionality of the filter holes, the ASF is optimised by dicing parts of the material from the middle, leaving three plates with holes separated by air gaps instead of continuous cylindrical holes (**Figure 4.1 a**).
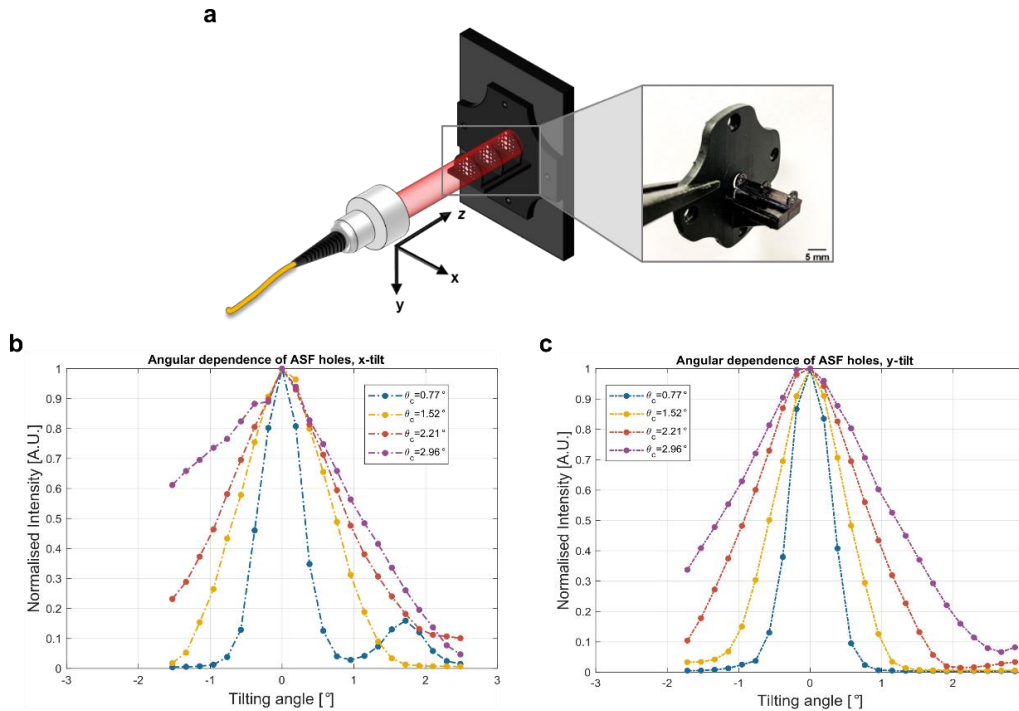
***Figure 4.1: Angular dependence of the ASF holes.** (a) The LED is tilted at various angles using a 3-axis translation stage. The x-y transverse plane is parallel to the CMOS sensor plane; the angular dependence*

*of the ASF holes is plotted as intensity vs. tilting angle in (**b**) x and (**c**) y direction. The full range of negative tilting angles could not be measured due to limitations in the maximum tilting angle allowed by the translation stage. The zoomed in picture of (**a**) shows the modified ASF.*

By reducing the thickness of each plate to 0.5 mm, the internal reflection is decreased significantly. The two plates on each end of the filter act like pinholes and the middle plate help in minimising crosstalk between the neighbouring holes. This new ASF has a length of 15 mm and consists of 19 holes, with filter cut-off angles ranging from 0.77° to 2.96°.

To test the angular dependence of the new ASF holes, the incident light source was gradually tilted with respect to the optical axis at different angles, first in the x and then in the y direction (**Figure 4.1 a**). An image was captured for each tilt angle, and the intensity in each filter hole was calculated. The changes in light intensity with varying tilt angles are shown in **Figure 4.1 b** and **c** for selected filter holes. It can be seen that there is an almost 90% decrease in light intensity at the cut-off angle for each filter hole shown. This demonstrates that the design efficiently reduces internal reflection and improves the angular dependence of the ASF holes.

The Mie scattering intensity is proportional to the square of the particle diameter[25]. Hence, to measure particles below 10 µm, it is necessary to improve the sensitivity of the PSA in measuring low scattering signals from small particles. To achieve this, the system is modified by using a pair of lenses and a beam stop placed at the focal point of the first lens to remove the unscattered forward light[26]. The experimental set-up is depicted in **Figure 4.2**. The ASF now only collects scattered light from the sample and hence is more sensitive to not only changes in scattered intensity caused by a change in sample concentration, but also to the shape and refractive index of the particle to be measured[27]. A typical raw image from the CMOS sensor is also shown in the figure.
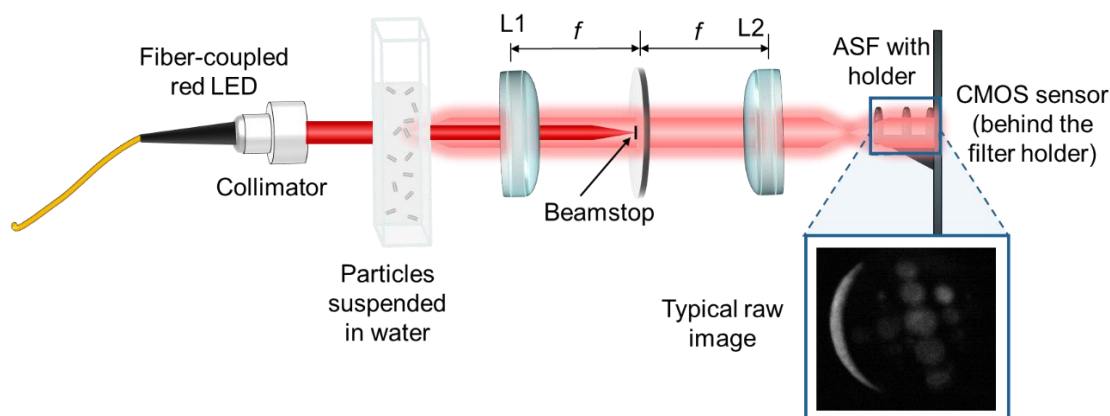
Figure 4.2: Modified PSA. *A schematic (not drawn to scale) of the modified PSA. L1 and L2 are anti-reflection (AR) coated biconvex lenses with equal focal length, f, of 50 mm. At the focal point of L1, a beam stop is placed by putting two layers of a small (approximately 1.2 mm square) piece of black tape on an AR-coated glass to block the unscattered light; a typical raw image from the CMOS sensor for a 2 µm particle at a concentration of 0.08 mg mL^{-1} is also shown.*

4.3 System characterisation with Silica Microspheres

In order to characterise the modified PSA, silica standard particles from Corpuscular Inc. with diameters of 2,5,10 and 15 µm at different concentrations were measured. Details of the silica beads are summarised in **Table 4.1**.

Table 4.1 Characteristics of measured silica microspheres

Sample	Refractive index	Density (g cm^{-3})	Mean size (µm)	Coefficient of Variation (COV) by manufacturer (%)	Concentrations tested (mg mL^{-1})
Silica beads	1.42	2.65	2	2	0.016, 0.08, 0.4, 2, 10, 50
			5	6	
			10	4.3	
			15	6	

Silica microspheres were chosen in this case as silica has a refractive index of 1.42, similar to that of *E. coli* and *Enterococcus sp.* For each sample, five 1:5 dilutions were prepared from the mother stock at a concentration of 50 mg mL^{-1} deionised water (DI), and measured in transparent, disposable plastic cuvettes with a light path length of 2 mm. Before adding the beam stop to block the forward light, a reference image was captured to determine the location of the filter holes and, subsequently, the light intensities in the holes were calculated. After the beam stop was placed, 15 background images were captured without any sample and later subtracted from the sample images. Prior to measuring each sample in the cuvette, it was shaken well to ensure a homogeneous dispersion of particles in the solution. This process was repeated after every 5 images in order to avoid particle sedimentation at the bottom, and a total of 15 images were acquired per sample. The images acquired in this study were captured with a CMOS image sensor, details of which were given in the previous chapter, and analysed using Matlab and Python.

The measurements were repeated four times in order to obtain four individual datasets. The average intensities calculated from all the datasets against concentration for each particle size are shown in **Figure 4.3**. As expected, the graphs show a monotonic increase in intensity up to a certain concetration for all particle sizes. The decrease in intensity for higher concentrations is a clear indication of a multiple scattering effect. Since multiple scattering effects are more prominent in smaller particles, the decrease in intensity is observed at a lower concentration for the 2 µm particle compared to the other sizes. The graphs of particle sizes 5 µm, 10 µm and 15 µm shows similar profile with increasing concentrations. This is probably due to an overlap in the particle size distribution between these sizes.
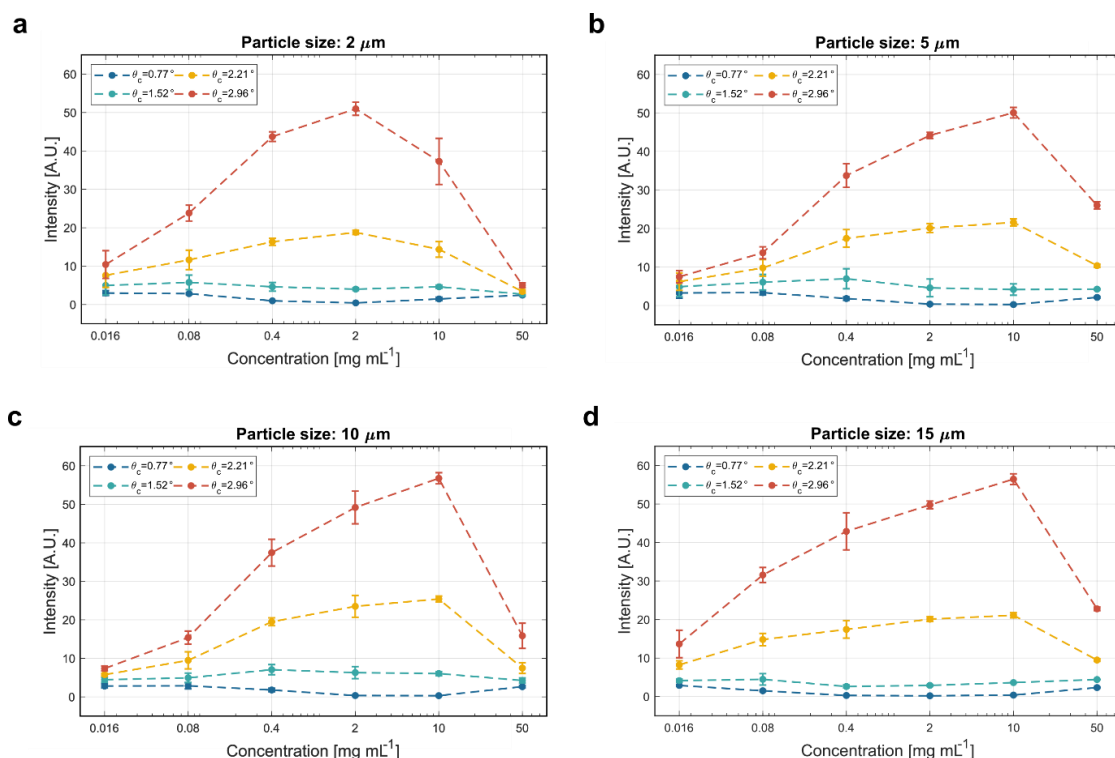
Figure 4.3: System characterisation using silica particles. *Average intensities of four selected filter holes at cut-off angles 0.77°, 1.52°, 2.21° and 2.96° for particle sizes (**a**) 2 µm, (**b**) 5 µm, (**c**) 10 µm and (**d**) 15 µm are plotted as a function of concentration. The error bars represent the standard deviation in the four data sets obtained for each particle size.*

4.4 Case-Study: Measurement of *E. coli* and *Enterococcus sp.* in Deionized Water

After characterising the PSA with microspheres, we performed measurements to determine whether the system can be used to detect different concentrations of *E. coli* and *Enterococcus sp.* in DI water. The *E. coli* strain used for all measurements was Op50[28]. To begin with, bacterial cultures were prepared by inoculating single isolated colonies of *E. coli* in a Luria-Bertani (LB) medium and grown at 37 °C in an incubator shaker (Thermo Fisher MaxQ8000). The bacterial growth was monitored by measuring the optical density (OD) at 600 nm using a spectrophotometer (Thermo Fisher Scientific Nanodrop 2000c). To prepare the samples, the culture media was collected at $OD_{600nm} \cong 1.5$ during the logarithmic growth phase of the bacteria when the concentration was ~10^9 cells mL^{-1} according to colony forming units (CFU). This was determined by plating serial dilutions of the suspension, in duplicate, on LB agar plates and counting the colonies after a 24 hour incubation at 37 °C. The culture media was centrifuged at 4,000 rpm and 27 °C for 5 minutes in order to pellet cells, after which the supernatant was removed and the cell pellets suspended in DI water. This process of centrifugation and re-suspension was repeated three times. The *E. coli* suspension was then used to prepare six tenfold serial dilutions in DI water and immediately measured with the PSA. The same procedure was followed to prepare *Enterococcus sp.* culture with the exception that the growth medium used was Tryptic Soy Broth (TSB) with yeast extract in this case. The culture media was collected at $OD_{600nm} \cong 1.6$ which corresponds to approximately 3.4×10^8 CFU mL^{-1}. Dilutions were prepared from this culture media in DI water and the final concentration of the samples was from ~10^3 to 10^8 cells mL^{-1}. The measurements were repeated four times on different days and four data sets were obtained.

The dependence of the scattering intensity on the concentration for *E. coli* and *Enterococcus sp.* is shown in **Figure 4.4** for DI water. Note here that the scattering intensity obtained from the PSA is from both viable and non-viable bacterial cells. In the case of *E. coli* (**Figure 4.4 a**) there is almost no intensity variation for concentrations from 10^4 to 10^6 cells mL^{-1} compared to the blank. For *Enterococcus sp.* (**Figure 4.4 b**) a little change in scattering intensity can be observed between the blank and concentrations 10^3 to 10^5 cells mL^{-1} beyond which the scattering signal increases monotonically up to about 10^7 cells mL^{-1}. The decrease in signal after this value is possibly due to the shadowing effect from the aggregation of particles in combination with the multiple scattering effect.
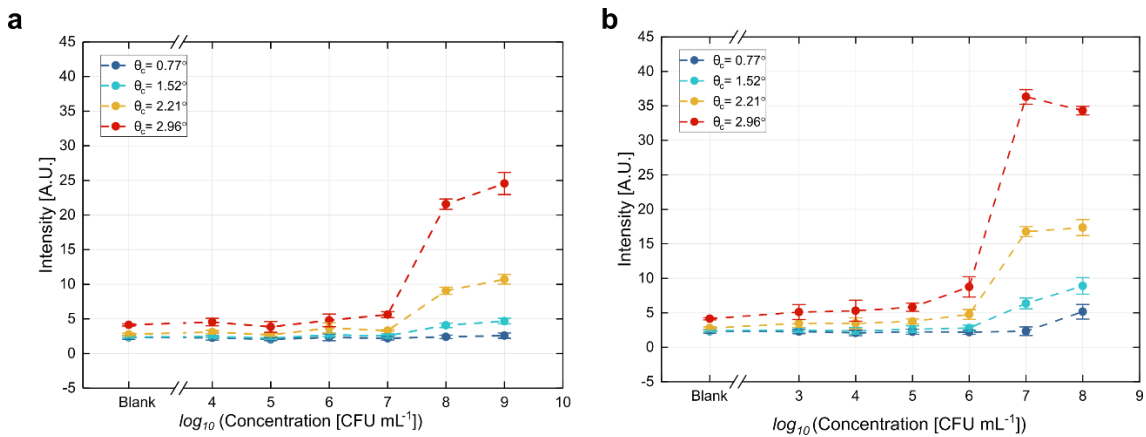
Figure 4.4: Measurement of E. coli. *Mean intensities for four ASF holes calculated from four measurement sets against log10 of concentration for (**a**) E. coli in DI water and (**b**) Enterococcus sp. in DI water. The error bars represent the standard deviation in the measurement sets.*

4.5 Data Analysis Using Machine Learning

In this work, we developed two ML models, one to classify between *E. coli*, *Enterococcus sp.* and particles in DI water, and the other to predict the concentration for a fixed sample. For both algorithms, we selected 9 holes out of 19 holes in the ASF images, based on the tilting experiment performed previously, discarding the holes that did not maintain the desired angular dependence due to geometrical or material imperfections. The datasets for all samples were split timewise, i.e., the most recent measurement set for each sample was used for testing and the rest for training the models. Both the models were trained with 100 decision trees for feature selection. The steps taken for the two models are summarised in the flowchart in **Figure 4.5** below.
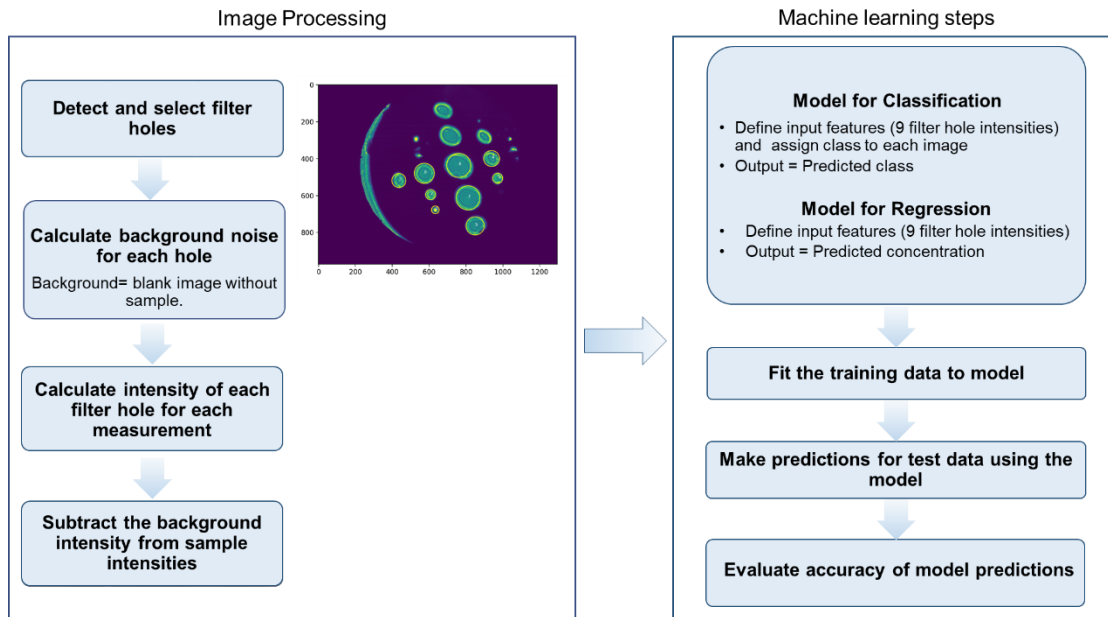
Figure 4.5: *Flow-chart showing the steps in the ML models used for classification and regression analysis.*

4.5.1   Classification between Microorganisms and Particles

The first model was developed to differentiate between the two types of bacteria and the four particles sizes in DI water and is based on Random Forest (RF) Classification[29]. This was the method of choice for our specific classification problem, mainly due to the short prediction time, higher accuracy and robustness with lower risk of overfitting[30].

The steps for the algorithm were essentially the same as those explained in the previous chapter, with the exception that it was a classification problem rather than one of regression. The first steps included locating the holes from the images, selecting the desired holes and calculating their intensities. Each sample image was assigned a numeric label provided in **Table 4.2** below.

Table 4.2: Label assigned to sample images for ML classification

Sample	Numeric label assigned
E. coli	0
Enterococcus sp.	1
2 µm	2
5 µm	5
10 µm	10
15 µm	15

The ML algorithm was then trained using the training set and predictions made on the test set, previously unseen by the algorithm. To visualise the performance of the ML model, a confusion matrix was plotted and the accuracy of the model was calculated as follows:

$$Accuracy = \frac{No.\,of\,correctly\,classified\,samples}{Total\,no.\,of\,samples} \times 100\,\%$$

Eq. 4.1

For the ML analysis, we selected the dataset based on a concentration range where the scattering intensity from the samples showed a monotonic trend only, i.e., from 0.016 mg mL^{-1} to 2 mg mL^{-1} for particles and from 10^6 to 10^9 CFU mL^{-1} for *E. coli* and from 10^4 to 10^7 CFU mL^{-1} for *Enterococcus sp.* in DI water. We selected four concentrations for all the samples within this range in order not to bias the classifier with varying sizes of dataset per sample.
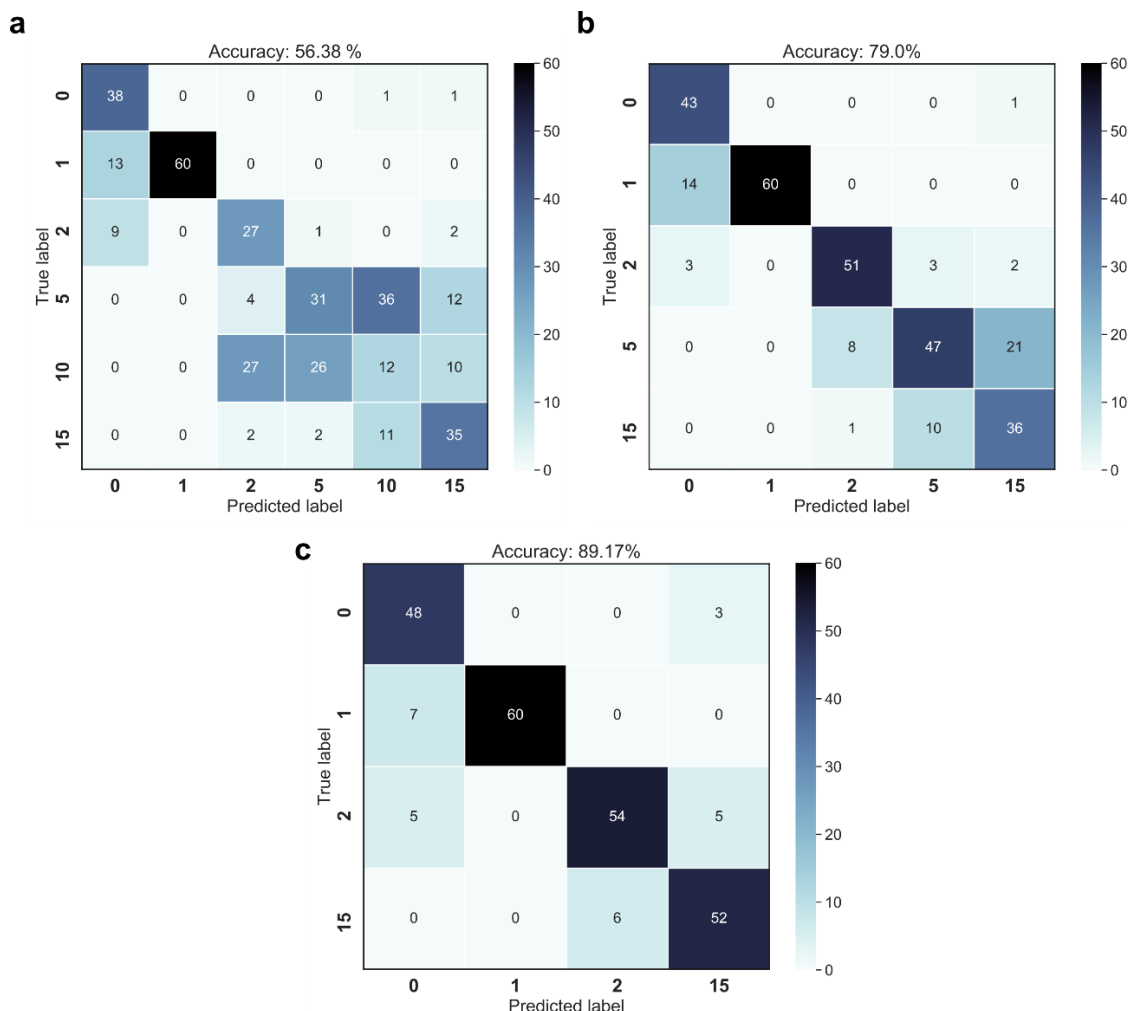
***Figure 4.6: Classification between E. coli, Enterococcus sp. and particles using RF classification model.** Confusion matrix for classification model (**a**) using all samples, (**b**) removing 10 µm and (**c**) removing both 5 and 10 µm particles. The samples corresponding to the numeric labels are listed in **Table 4.2**.*

The confusion matrix with all samples is depicted in **Figure 4.6 a**, demonstrating that the model can classify the samples with an accuracy of 56%. Interestingly, the model can classify very well *Enterococcus sp.* from the rest of the samples and as seen from the confusion matrix, all 60 images were predicted correctly by the RF classification model. A closer look at the confusion matrix reveals that the model cannot properly classify between different sizes of silica beads which is the reason for the low overall accuracy. Hence, we performed another ML analysis, removing the 10 µm particle (**Figure 4.6 b**) first in which case the accuracy improved to 79% and then both 5 µm and 10 µm particles (**Figure 4.6 c**) leading to accuracy of nearly 89%. These improvements confirm our hypothesis of overlapping size distributions mentioned before. In future, by collecting more data with well separated particle size distributions, the model may be improved.

To demonstrate that the multiple holes of the ASF play a crucial role in classifying the various particles and microorganisms, we performed the next analysis using only the biggest ASF hole.

The accuracy reduces dramatically from 89% to 49% as seen in the confusion matrix (**Figure 4.7 a**), indicating that absorption information from the particle solution alone is insufficient for the RF classifier to discriminate between different samples. Following that, we repeated the analysis, each time increasing the number of holes used by the model for classifying the samples. The holes were randomly selected and for a fixed number of holes, the analysis was performed three times with different combinations of hole sizes. In **Figure 4.7 b**, the classification accuracy obtained is plotted against the number of holes. It can be seen that with a single hole, i.e. the big hole, the accuracy is 49% and with increasing number of holes, the accuracy also increases till it reaches a maximum at 10 holes, after which the accuracy seems to saturate at roughly 85% with some fluctuations.
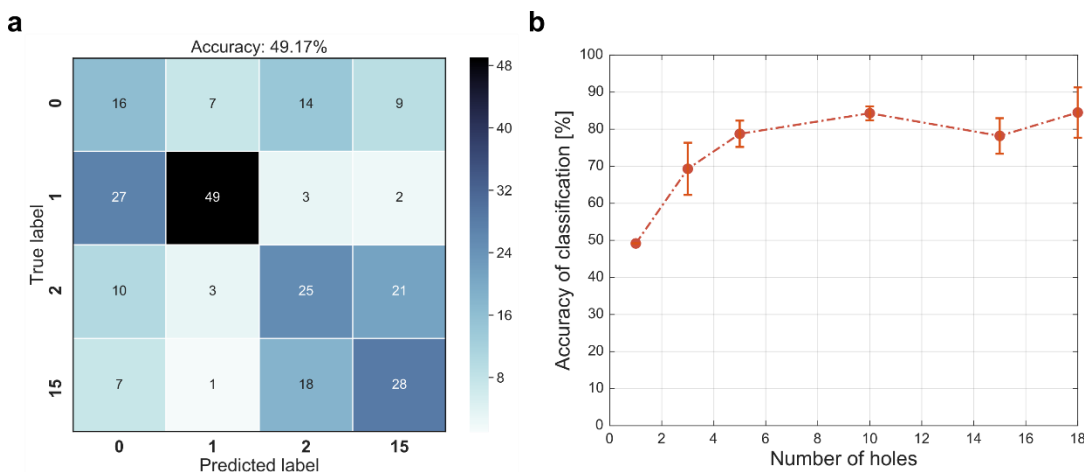
***Figure 4.7: RF classification analysis to demonstrate the functionality of the ASF holes.** (a) Confusion matrix using only the biggest ASF hole for classifying E. coli, Enterococcus sp., 2 µm and 15 µm. (b) Accuracy against number of holes. The analysis was repeated three times with different combination of hole sizes for a fixed number of holes. The graph shows average classification accuracy with error bars representing the standard deviation between the three analyses performed.*

4.5.2 Concentration Prediction for Particles and Microorganisms

After classifying the different samples, we developed another ML model using RF Regression[31] in order to predict the concentration for a fixed particle size. Intensities in the ASF holes were used to train the ML model and predictions were made on the test set. In addition, concentration ranges were selected using the aforementioned criteria. However, for particles greater than 2 µm we included the entire concentration range, i.e., 0.016 mg mL^{-1} to 10 mg mL^{-1}, where the graphs show an increasing trend. The metric used to evaluate the performance of the model was Mean Absolute Error (MAE). For a given test dataset with *n* number of images, the MAE is given by the following equation:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Actual\ value_i - Predicted\ value_i| \qquad \text{Eq. 4.2}$$

The smaller the MAE, the better and more stable the predictions obtained from the model. **Figure 4.8** depicts the mean predicted concentration for different particle sizes with respect to the corresponding nominal concentration measured for each particle size. The MAE, Pearson's R and R^2 values obtained from the prediction plots are reported in **Table 4.3** below.

Table 4.3: MAE, Pearson's R and R² for different particle sizes.

Sample	MAE	Pearson's R	R^2
2 µm	0.15	0.99	0.99
5 µm	0.79	0.97	0.92
10 µm	2.12	0.78	0.30
15 µm	0.4	0.99	0.98

It can be observed from the graphs that the mean predictions for 2 µm, 5 µm and 15 µm correlate well with the actual concentration, whereas for 10 µm, the predictions show a significant deviation from the nominal concentrations.
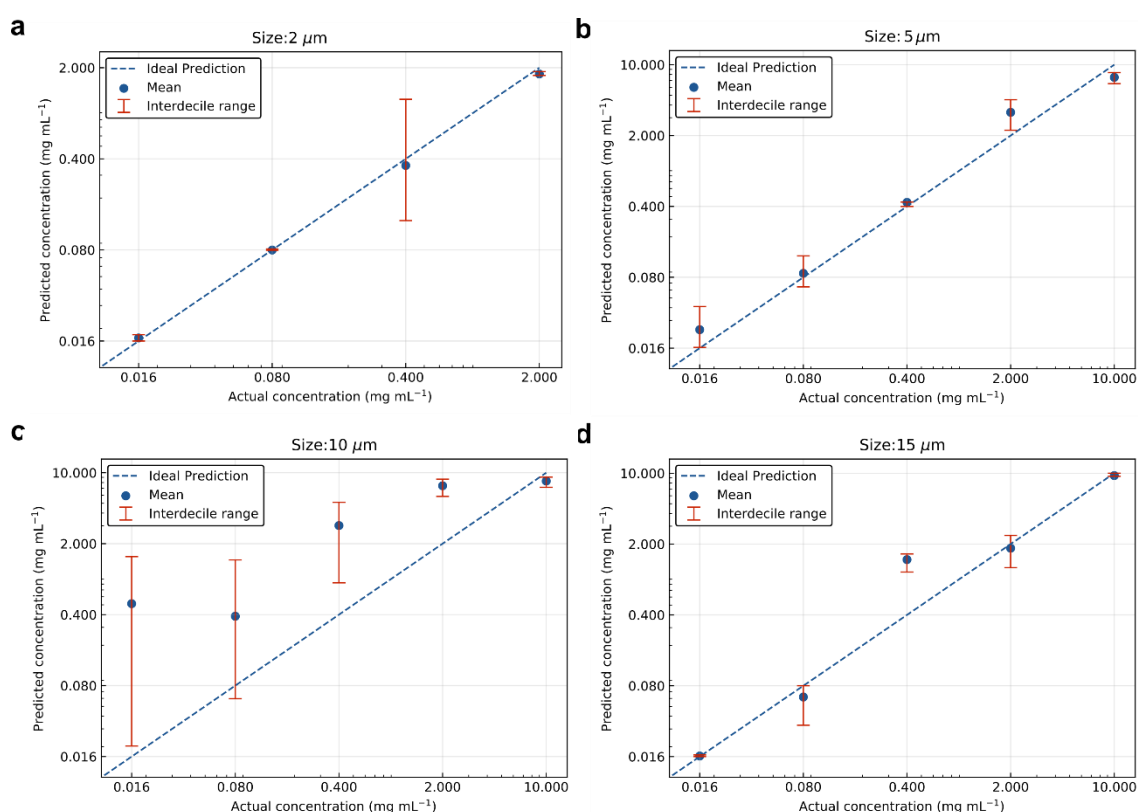
Figure 4.8: Predicted concentration of particles using RF regression model. *The mean predicted concentration values for test set 4 plotted against nominal concentration for (**a**) 2 µm, (**b**) 5 µm, (**c**) 10 µm and (**d**) 15 µm. The dashed lines represent predicted concentration = actual concentration. The interdecile range is also shown for each concentration.*

The graphs for the same analysis performed with *E. coli* and *Enterococcus sp.* in DI water are plotted in Error! Reference source not found. **a** and **b**, respectively, showing good agreement between the predicted and measured concentrations. The corresponding MAE are shown in **Table 4.4**. The R^2 score and the Pearson correlation coefficient, R, obtained for the mean predictions fitted on the measured log10 CFU mL^{-1} (**Table 4.4**), further confirm a strong correlation between actual and predicted concentrations.
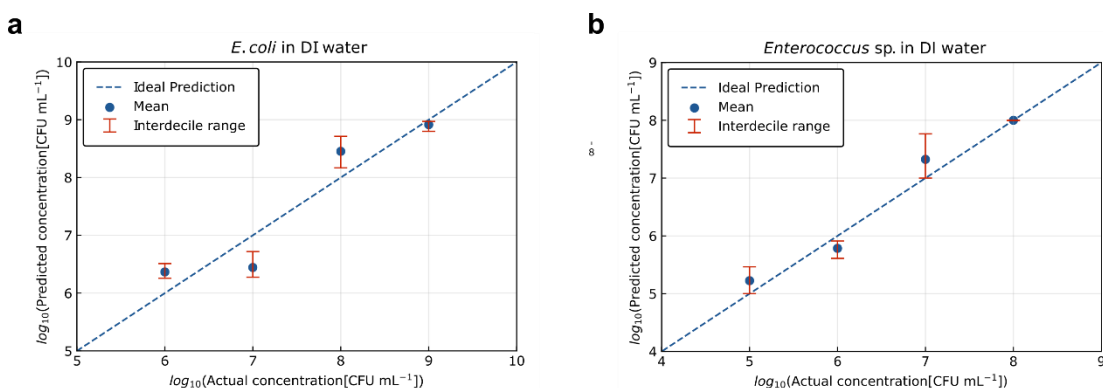
Figure 4.9: Predicted E. coli concentrations using RF regression model. *Mean concentration prediction on test set 4 for (**a**) E. coli in DI water, MAE = 0.36 and (**b**) Enterococcus sp. in DI water, MAE= 0.16. The dashed line represents y=x and the interdecile range is shown for each concentration.*

Table 4.4: MAE, Pearson correlation (R) and R^2 score on the mean predictions for the test dataset

Sample	MAE	Pearson's R	R^2
E. coli in DI water	0.36	0.94	0.88
Enterococcus sp. in DI water	0.16	0.98	0.97

The ML model can be improved with further optimisations to the system and by increasing the data size used for training the model. The major benefit of the PSA is the incredibly short (< 10 minutes) time-to-result, and, although much development is needed to improve the sensitivity of the system, the PSA has the potential to be used as a rapid preliminary screening tool for bacterial contamination.

4.6 Conclusion

In this chapter, we demonstrated the potential of our PSA for biological applications such as the detection and classification of *E. coli* concentration in water from 10^6 to 10^9 CFU mL^{-1} and *Enterococcus sp*. from 10^5 to 10^8 CFU mL^{-1}. In order to achieve this, we first modified the system to measure small particles (below 20 µm) by removing the forward unscattered light, and characterised it by measuring several concentrations of silica microspheres with similar particle sizes. The results indicated that the modified PSA is, in fact, able to detect smaller particles and is more sensitive to concentrations lower than those measured with the previous system. To investigate the detection limit of the system for microorganism measurement, we tested two types of microorganisms, *E. coli* and *Enterococcus sp*. suspensions in DI with concentrations ranging from ~10^4 to 10^9 cells mL^{-1} and from ~10^3 to 10^8 cells mL^{-1} in the former and latter case, respectively. The initial results revealed an increasing signal trend for *E. coli* starting at 10^6 cells mL^{-1} and 10^5 cells mL^{-1} for *Enterococcus sp*. Below these concentrations, the low scattering signal resulted in a poor signal-to-noise ratio, and the system was unable to distinguish bacterial suspensions presence. Next, we developed a ML algorithm to distinguish *E. coli* from particles in DI water. The initial model accuracy obtained was 56% when trained with data within the relevant concentration range of the PSA. The reason for this low accuracy was mainly due to overlapping size distributions for particles above 2 µm. When 10 µm and then both 5 µm and 10 µm particles were removed from the analysis, the accuracy improved to 72% and 89%, respectively. In both cases, we observed that the model could classify well between *E. coli* and *Enterococcus sp*. with

only few false predictions, even though their mean sizes are very similar. This is encouraging as it suggests that the PSA is capable of differentiating between similar sized particles with different shapes and refractive indices. We also demonstrated that, if the classification analysis was performed with only the largest hole, which contains mostly the absorption information about the measured solution, the accuracy dropped significantly. This clearly shows that a single hole is insufficient, and that the ASF's multiple holes are essential for classifying different samples.

On successfully distinguishing between the two types of bacteria and particles, we developed another model to predict the concentration for a fixed sample. The results obtained for the particles and the bacteria showed a strong correlation between the actual and the predicted concentrations, except for 10 µm particles. For *E. coli* and *Enterococcus sp.* in DI water, we obtained good predictions with a low MAE for a concentration range from ~10^6 to 10^9 cells mL^{-1} and ~10^5 to 10^8 cells mL^{-1}, respectively. In the future, the detection limit of the PSA may be improved by further modifying the system, for example, by combining a blue LED to enhance the scattering signal from bacterial cells with a more sensitive CMOS sensor and an ASF filter with bigger holes or a shorter length in order to efficiently collect the large scattering angles. With these developments, the PSA could become a rapid, easy to use, low-cost on-site monitoring tool to determine levels of particulates, for example, downstream of a waste water treatment plant, as an increased particle concentration is an indication of treatment failure.

One current limitation of the system is its lack of specificity as a result of its true label-free detection nature, which prevents its widespread application. The PSA could be made more specific in the future by using capture mechanisms like antibody-coated magnetic particles to bind to target bacteria, which could then be separated from other particles in liquid suspensions. This would enable rapid identification of bacteria in food or other relevant samples without the need for colonies to be grown or overnight cultures to be prepared, which takes 12-16 hours on average. However, using light scattering to detect particle sizes in the nanometre range will likely remain a challenge for the proposed PSA system. Hence, the focus of the next chapter will be to design a small form factor optical biosensor based on single particle signal, a flow-cytometry which uses fluorescence for rapid, sensitive and specific detection of particles smaller than bacterial cells, specifically viruses.

4.7 Reference

1.   Rajapaksha, P. *et al.* A review of methods for the detection of pathogenic microorganisms. *Analyst* **144**, 396–411 (2019).

2.   Li, J., Zhu, Y., Wu, X. & Hoffmann, M. R. Rapid Detection Methods for Bacterial Pathogens in Ambient Waters at the Point of Sample Collection: A Brief Review. *Clin. Infect. Dis.* **71**, S84–S90 (2020).

3.   Lazcka, O., Campo, F. J. Del & Muñoz, F. X. Pathogen detection: A perspective of traditional methods and biosensors. *Biosens. Bioelectron.* **22**, 1205–1217 (2007).

4.   Chatterjee, A. & Abraham, J. Chapter 2 - Microbial Contamination, Prevention, and Early Detection in Food Industry. in *Microbial Contamination and Food Degradation* (eds. Holban, A. M. & Grumezescu, A. M.) 21–47 (Academic Press, 2018).

5.   De Boer, E. & Beumer, R. R. Methodology for detection and typing of foodborne microorganisms. *Int. J. Food Microbiol.* **50**, 119–130 (1999).

6.   Yasir, M. Analysis of Microbial Communities and Pathogen Detection in Domestic Sewage Using Metagenomic Sequencing. *Divers. 2021, Vol. 13, Page 6* **13**, 6 (2020).

7.   Ajonina, C., Buzie, C., Rubiandini, R. H. & Otterpohl, R. Microbial pathogens in Wastewater Treatment Plants (WWTP) in Hamburg. *J. Toxicol. Environ. Heal. - Part A*

Curr. Issues **78**, 381–387 (2015).

8. Noble, R. T. & Weisberg, S. B. A review of technologies for rapid detection of bacteria in recreational waters.

9. Law, J. W.-F., Mutalib, N.-S. A., Chan, K.-G. & Lee, L.-H. Rapid methods for the detection of foodborne bacterial pathogens: principles, applications, advantages and limitations. *Front. Microbiol.* **5**, (2014).

10. Valones, M. A. A. *et al.* Principles and applications of polymerase chain reaction in medical diagnostic fields: a review. *Brazilian J. Microbiol.* **40**, 1–11 (2009).

11. McCarthy, J. Immunological techniques: ELISA. in *Detecting Pathogens in Food* **84**, 241–258 (Elsevier, 2003).

12. Hussain, M. *et al.* Rapid Identification of Pathogens based on MIE Light Scattering and Machine Learning Approach. *Med. Meas. Appl. MeMeA 2019 - Symp. Proc.* (2019). doi:10.1109/MEMEA.2019.8802228

13. Liang, P. S., San Park, T. & Yoon, J.-Y. Light scattering-based detection of food pathogens. in *Light Scattering Technology for Food Property, Quality and Safety Assessment* 429–444 (CRC Press, 2017).

14. Bhunia, A. K., Banada, P., Banerjee, P., Valadez, A. & Hirleman, E. D. Light scattering, fiber optic- and cell-based sensors for sensitive detection of foodborne pathogens. *J. Rapid Methods Autom. Microbiol.* **15**, 121–145 (2007).

15. Sang, S. *et al.* Progress of new label-free techniques for biosensors: a review. *Crit. Rev. Biotechnol.* **36**, 1–17 (2015).

16. O, L., FJ, D. C. & FX, M. Pathogen detection: a perspective of traditional methods and biosensors. *Biosens. Bioelectron.* **22**, 1205–1217 (2007).

17. Cooper, M. A. Label-Free Biosensors: Techniques and Applications. *Label-Free Biosens. Tech. Appl.* 1–304 (2009).

18. Wieliczko, A., Podbielska, H. & Buzalewicz, I. Influence of various growth conditions on Fresnel diffraction patterns of bacteria colonies examined in the optical system with converging spherical wave illumination. *Opt. Express, Vol. 19, Issue 22, pp. 21768-21785* **19**, 21768–21785 (2011).

19. Abdelhaseib, M. U., Singh, A. K. & Bhunia, A. K. Simultaneous detection of Salmonella enterica, Escherichia coli and Listeria monocytogenes in food using a light scattering sensor. *J. Appl. Microbiol.* **126**, 1496–1507 (2019).

20. Haavig, D. L. *et al.* Cellular Light Scattering for the Identification of Bacteria and Its Application to the Identification of Staphylococcus. *J. AOAC Int.* **100**, 1836–1847 (2017).

21. Molleda, P., Blanco, I., Ansola, G. & de Luis, E. Removal of wastewater pathogen indicators in a constructed wetland in Leon, Spain. *Ecol. Eng.* **33**, 252–257 (2008).

22. García-Solache, M. & Rice, L. B. The enterococcus: A model of adaptability to its environment. *Clinical Microbiology Reviews* **32**, (2019).

23. Rompré, A., Servais, P., Baudart, J., De-Roubin, M. R. & Laurent, P. Detection and enumeration of coliforms in drinking water: Current methods and emerging approaches. *J. Microbiol. Methods* **49**, 31–54 (2002).

24. Subgingival Microbes. in *Atlas of Oral Microbiology* 67–93 (Elsevier, 2015). doi:10.1016/b978-0-12-802234-4.00004-5

25. Gotoh, K., Masuda, H. & Higashitani, K. Powder technology handbook. *Journal of Thermal Spray Technology* **10**, 10 (2001).

26.   Smith, Z. J., Chu, K. & Wachsmann-Hogiu, S. Nanometer-Scale Sizing Accuracy of Particle Suspensions on an Unmodified Cell Phone Using Elastic Light Scattering. *PLoS One* **7**, e46030 (2012).

27.   Jones, A. R. Light scattering for particle characterization. *Progress in Energy and Combustion Science* **25**, 1–53 (1999).

28.   Neve, I. A. A. *et al.* Escherichia coli Metabolite Profiling Leads to the Development of an RNA Interference Strain for Caenorhabditis elegans. *G3 Genes|Genomes|Genetics* **10**, 189–198 (2020).

29.   Liaw, A. & Wiener, M. Classification and Regression by randomForest. **2**, (2002).

30.   Amaratunga, D., Cabrera, J. & Lee, Y.-S. Enriched random forests. *Bioinformatics* **24**, 2010–2014 (2008).

31.   Breiman, L. Random Forests. *Mach. Learn. 2001 451* **45**, 5–32 (2001).

CHAPTER 5

Portable Flow-Virometer for Detection of SARS-CoV-2 in Saliva

Under the rules of the Creative Commons Attribution-NonCommercial License, the content, text, and figures in this chapter have been adapted from the original publication: "Small form factor flow virometer for SARS-CoV-2", Hussain, R.*, Ongaro, A.*, et al. Biomed. Opt. Express 13, 1609-1619 (2022).

Author contribution: In this work I contributed to designing the experiments, building the FVR prototype, preparing and measuring the samples. I performed the STED imaging together with another co-author of the paper and wrote the code for the preliminary analysis from the raw data. In addition, I co-wrote the first draft of the manuscript.

5.1 Introduction

Throughout human history, some of the deadliest pandemics have been caused by sub-microscopic viruses, often known as ¨the organisms at the edge of life¨[1]. The most recent life-threatening pandemic witnessed by the world is the Coronavirus Disease 2019 or COVID-19, caused by Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2), infecting over 160 million people worldwide[2]. Such a viral outbreak as COVID-19 made evident how critical is detecting and isolating infected patients quickly and effectively. With regards to this, a low-cost, fast, accurate, and portable diagnosis system is potentially a breakthrough technology for controlling the pandemic, returning to normal life quickly and reducing the risk of future outbreaks[3–5].

SARS-CoV-2 is a single-stranded RNA-enveloped virus[6] and its full genome sequencing has enabled the development of various diagnostics. A reverse transcription polymerase chain reaction (RT-PCR), which has been endorsed by the Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO), is the current gold standard testing system for diagnosing suspected COVID-19 casesc. Reverse transcription loop-mediated isothermal amplification (RT-LAMP)[11] is a more recent breakthrough in molecular testing. Although RT-LAMP assays, unlike RT-PCR, do not require expert staff and are, thus, well suited for point-of-care (POC) and self-testing[12], achieving a quantifiable response with this method is still challenging. Both of these approaches involve RNA extraction, which is time-consuming and requires the use of resource-intensive processes, thus increasing the overall cost.

Antigen tests, on the other hand, can detect easy-to-find surface markers (viral antigens) on the outside of the virus, and do not require extraction and amplification steps[13–15]. SARS-CoV-2 has two known antigens, nucleocapsid phosphoproteins and spike glycoproteins, capable of attacking cells in the human body. The main principle behind antigen detection tests is the use of designed, synthetic antibodies to probe a patient sample for the presence of viral proteins, which can be detected in the blood or in other tissues or secretions, such as saliva. While there are some shortcomings to the sensitivity and specificity of this method compared with RT-PCR or RT-LAMP methods, it is much easier, cheaper and faster, and can be performed anywhere[15–21].

Apart from molecular assays, optical biosensors are a promising alternative method for diagnosing SARS-CoV-2 because of their high sensitivity, ease of use, quick time-to-result, and point-of-care testing capabilities. Label-free optical sensors based on surface plasmon resonance (SPR), localised surface resonance (LSPR), and surface-enhanced Raman spectroscopy (SERS) have recently received a lot of attention[22]. Other new optical approaches for detecting SARS-CoV-2 include spatial light-interference microscopy (SLIM) in combination with neural network or nano-interferometric biosensor[23,24]. Although preliminary results reported in the literature suggest that these technologies have the potential to be employed as a highly sensitive SARS-CoV-2 detection platform, there is still a lot of work to be done before they can be used outside of the lab and introduced into the commercial phase.

Recent evidence shows that saliva is a more viable and sensitive alternative to nasal and throat swabs[17,25]. The use of saliva samples has a number of clinical benefits: (1) it is non-invasive and improves the patient experience, increasing voluntary testing take-up; (2) the collection can be self-administered with simple instructions, reducing the risk of viral transmission to healthcare personnel; and (3) it can be used to quickly check for the presence of the virus in asymptomatic individuals[26]. Therefore, the use of saliva allows for large-scale SARS-CoV-2 testing, particularly when combined with a rapid and sensitive detection approach[27].

Flow-cytometry[28] has been used for decades as a powerful laser-based cell analysis technology for counting and sorting cells based on light scattering and/or fluorescence. However, conventional flow-cytometers are bulky (~ 42.5×42.5×34 cm) devices with prices ranging from 30,000€ to 300,000€, making them unsuitable for POC applications. In recent years, flow-cytometry has paved the way for the development of flow-virometry[29], a technique for the direct and accurate detection and characterisation of nanoparticles like viruses, which can be achieved by labelling viral particles with fluorescently tagged antibodies. For this purpose, simple and compact flow-virometers for specialised applications can be designed that are easy to use and have minimal overall costs. Therefore, the main objective of this chapter is to introduce a cost-effective and portable (25×30×13 cm) flow-virometry reader (FVR) for the rapid detection of whole SARS-CoV-2 viral particles directly from saliva samples, without genetic extraction, thus combining the convenience and simplicity of an antigen test with the high precision of RT-PCR. We will explain the development and optimisation processes of the FVR assay in depth, as well as how the system is validated using clinical data. Finally, we will report the limit of detection (LOD), sensitivity and specificity of the developed sensor platform.

5.2 Concept of the FVR for SARS-CoV-2 Detection from Saliva

When it comes to finding faster, simpler, and more accurate SARS-CoV-2 detection procedures, combining precise flow virometry with a quick immunoassay test can increase the speed and precision of SARS-CoV-2 detection from saliva. In **Figure 5.1,** we demonstrate the entire workflow of the developed diagnostic system, from the sample collection through to the SARS-

CoV-2 antigen detection. In brief, saliva is collected in sterile tubes through spitting, with no restrictions on when food or drink has previously been consumed. It is then purified using a simple sample preparation technique based on heat inactivation at 56 °C for 1 hour, followed by dilution and filtering. After that, the saliva is labelled with fluorescent anti-SARS-CoV spike antibodies and incubated at room temperature for 20 minutes. The labelled sample is circulated inside our developed FVR and a fluorescence signal is detected every time the sample passes through the interrogation point. We note that in some cases detection of particles can be performed without fluorescence labels, as it has been the case of work presented in chapters 3 and 4. However, in the case of viruses, given their small size of the order of 100 nm, labels are needed to increase detection signal and achieve in this way high sensitivity. In addition, labels with proper antibody or ribonucleic acid (RNA) allow reaching specificity.
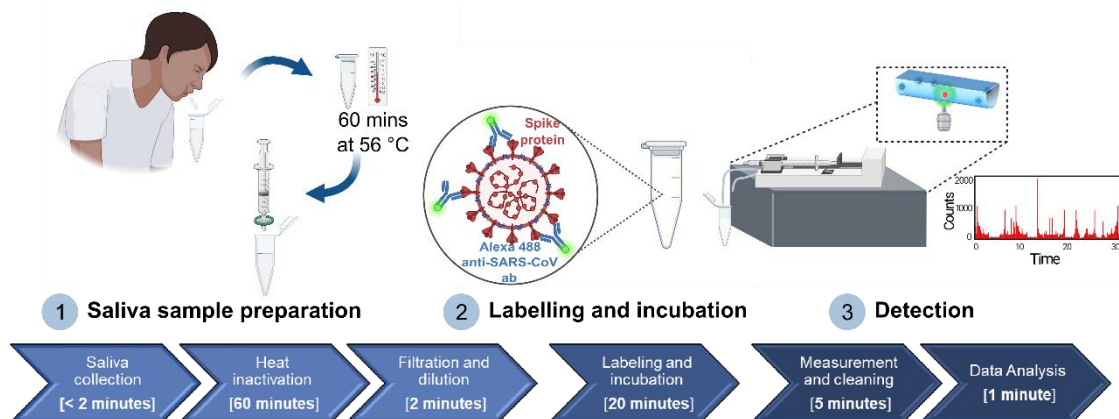
Figure 5.1: A comprehensive overview of the developed diagnostic system used for the detection of SARS-CoV-2 antigens.

5.3 Development and Characterisation of the FVR

The proposed FVR is a compact, portable flow-virometer, which does not require flow focusing, and which is tailored for the detection and quantification of only the fluorescence signal from the labelled SARS-CoV-2 viral particles. **Figure 5.2 a** depicts the schematic diagram of the developed FVR, and a lab-built prototype with the tablet to control the whole system is shown in **Figure 5.2 b.** An excitation beam from a fibre-coupled 488 nm laser is focused onto the microfluidic channel by a 100x microscope objective. As the sample passes through the laser intercept, the fluorophores in the sample emit fluorescence, which is then directed to a multi-pixel photon counting (MPPC) detector via a dichroic mirror and two emission filters. The optical filters ensure that only the emitted fluorescence signal is detected, without any stray light from the excitation source, thus reducing the background noise and improving the system's sensitivity.

To calibrate the FVR, a known concentration of fluorescent polystyrene beads (with a nominal diameter of 2 µm) in phosphate-buffered saline (PBS) was measured. In order to ensure proper alignment of the optical system with respect to the straight channel microfluidic chip, the bead solution, with a concentration of 250 beads mL^{-1}, was circulated inside the microfluidic channel at a flow rate of 1 mL min^{-1}. The position of the chip was then adjusted until a steady rate of at least 2 events per second was observed. For the calibration measurements, five 10-fold dilutions of the polystyrene bead solution were measured in triplicate. **Figure 5.2 c** shows the calibration plot obtained for the polystyrene particles, with concentrations ranging from 2.5 beads mL^{-1} to 2.5×10^4 beads mL^{-1}. Over the concentration range tested, the graph depicts a sharp linear response ($R^2 = 0.98$), showing that the total number of events counted in a 1 mL sample and the bead concentration in the solution are in good agreement.

In order to detect SARS-CoV-2 viral particles using the FVR, a baseline with the fluorescent antibodies must first be established. Four 10-fold serial dilutions, with concentrations ranging from 0.5 ng mL^{-1} to 500 ng mL^{-1}, were measured in triplicate to obtain the analytical response of the fluorescent (Alexa-488) anti-SARS-CoV spike antibodies. The results obtained from these measurements are plotted in **Figure 5.2 d**, showing a strong correlation between the prepared anti-SARS-CoV spike antibody concentrations and the detected fluorescent events ($R^2 = 0.94$; *p* = 0.001).
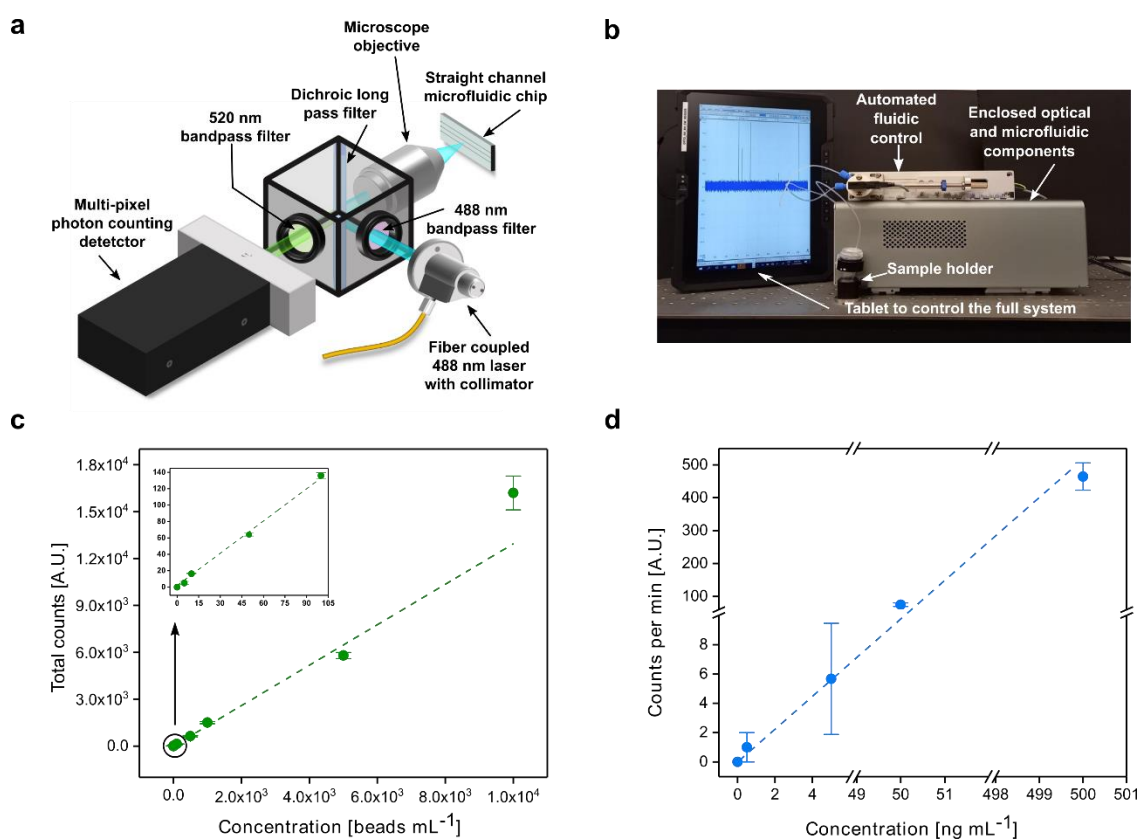
Figure 5.2: Design and characterisation of the FVR reader. (a) FVR detection system depicted in a schematic diagram; (b) a lab-built FVR prototype with the tablet; (c) standard curve of the polystyrene bead concentration vs. total counts, $R^2 = 0.99$; (d) standard curve of the fluorescent (Alexa 488) anti-SARS-CoV spike antibody concentration vs. counts per minute, $R^2 = 0.94$.

5.4 FVR Assay Optimisation for SARS-CoV-2 Antigen Detection

The FVR SARS-CoV-2 detection protocol relies on effective binding between the antibodies and the target antigen. Several parameters can influence the rate of this binding, including composition of the saliva samples, antibody-antigen concentration and temperature, amongst others[30]. As a result, these parameters were optimised to ensure proper formation of the antibody-antigen complex. Prior to sample preparation, the saliva samples were heat inactivated for 1 hour at 56 °C, due to the restrictions of performing the measurements in a Biosafety Level 3 facility[31]. All saliva samples collected from both infected and uninfected healthy donors were confirmed with an RT-qPCR test.

5.4.1 Saliva Pretreatment

Saliva is a complex fluid comprising of many substances, such as mucin, food debris, microorganisms and gingival crevicular fluid, and it differs widely from one individual to another.
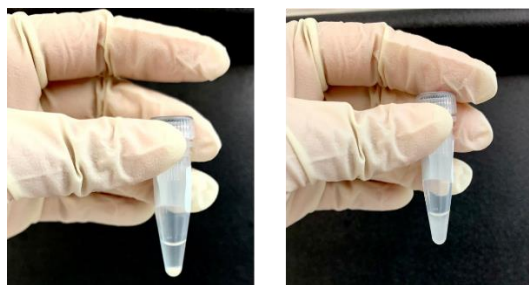
Figure 5.3: Picture of collected saliva as received (left) and after vortexing to re-suspend particles and heavy molecules (right).

To obtain a precise analysis, it is important to remove these impurities from the saliva without affecting the relevant analytes. However, for POC applications, a quick and simple pre-treatment method for the saliva is required, without the need for laboratory equipment such as centrifuges or freezers. In order to achieve this, we re-suspended the particles and heavy molecules by means of vortexing, and then used a 0.2 µm pore syringe filter to remove any large particles that would otherwise increase the background noise and cause blockages in the microfluidic channel and tubing in the FVR. **Figure 4.4** shows the heavy sediments of salivary contaminants before and after vortexing. The graphs in **Figure 5.4 a** and **b** demonstrate that the filtration process reduces the background noise from the saliva by 98%. This was also confirmed by the statistical analysis performed on the filtered and unfiltered saliva samples (**Figure 5.4 c**).
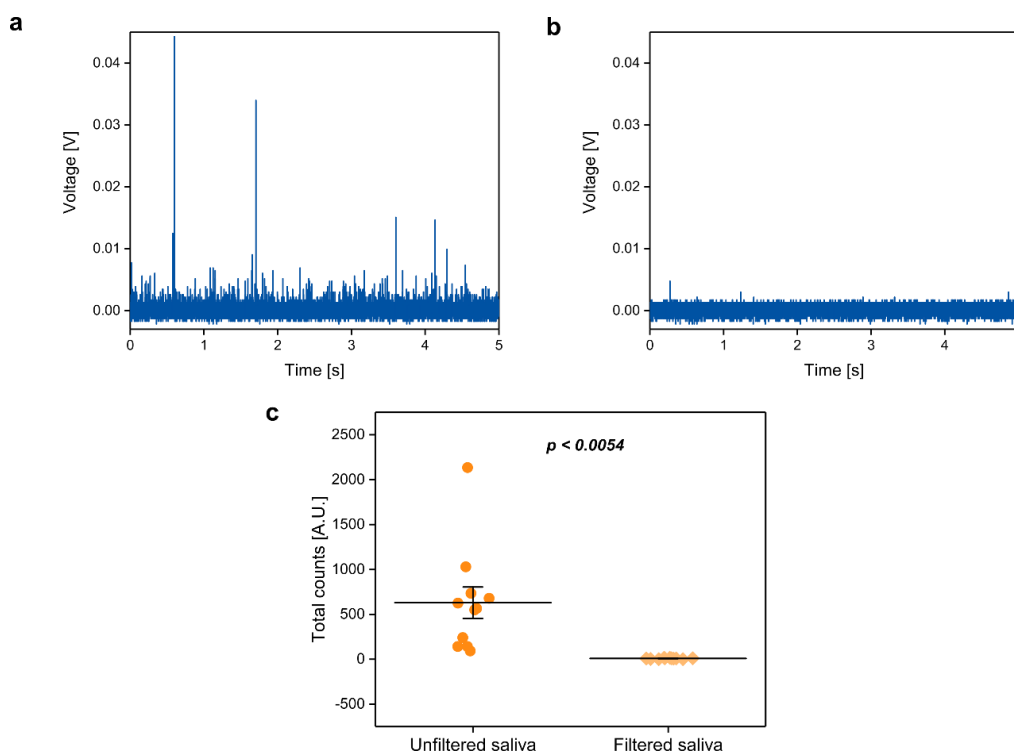
Figure 5.4: Background noise associated with the saliva samples. *Raw Voltage vs. time signal from oscilloscope of (**a**) unfiltered saliva sample and (**b**) filtered saliva sample; (**c**) difference in total counts*

between the unfiltered (left) and filtered (right) population of 11 saliva samples. The two sample populations are statistically different, with p < 0.0054 calculated with a two-tailed student t-test.

After filtering the saliva, it was then diluted 100-fold in PBS. These filtration and dilution steps help to reduce variations in the salivary viscosity and allow for a constant laminar flow in the microfluidic channel without the need for high pressure (> 13 bar).

5.4.2 Optimum Antibody Concentration and Incubation Time

The extent of antibody-antigen effective binding can also be controlled by varying the relative concentrations of an antibody and an antigen solution. Thus, for our assay, the optimal antibody concentration that provides the best labelling was determined experimentally. A series of 1:10 dilutions of a SARS-CoV-2 positive saliva sample (viral load of 10^7 copies mL^{-1}) was performed on healthy donor saliva, i.e., SARS-CoV-2 negative by RT-qPCR. The dilutions had final concentrations ranging from 10^7 copies mL^{-1} to 100 copies mL^{-1}. Each of these prepared dilutions was then labelled with four different Alexa-488 anti-SARS-CoV spike antibody concentrations (25 μg mL^{-1}, 5 μg mL^{-1}, 500 ng mL^{-1} and 50 ng mL^{-1}) and measured.
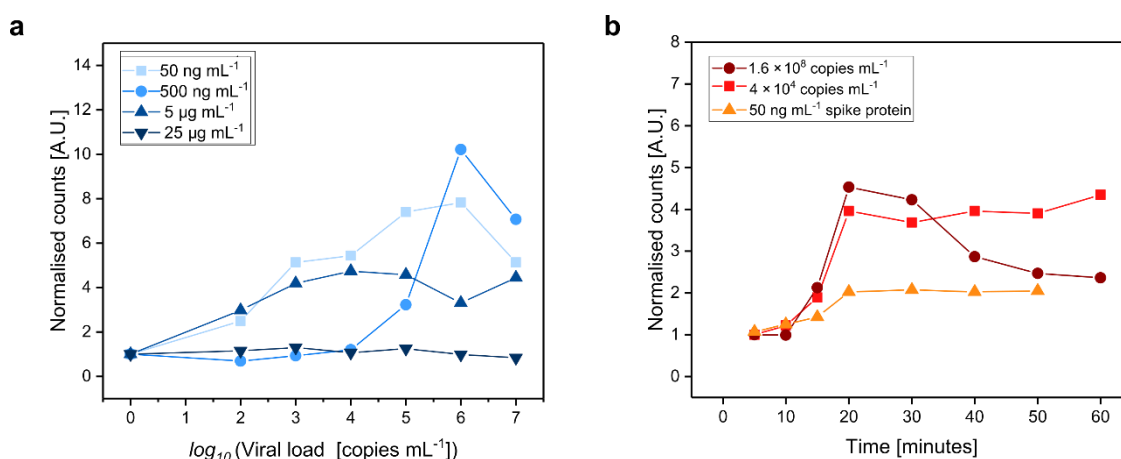
Figure 5.5: Labelling protocol optimisation. *(a) Normalised counts against log of viral load for four antibody concentrations, 25 μg mL^{-1}, 5 μg mL^{-1}, 500 ng mL^{-1} and 50 ng mL^{-1}; (b) Normalised counts from labelled saliva samples with 50 ng mL^{-1} anti-SARS-CoV spike antibodies against incubation time for two positive SARS-CoV-2 saliva samples and one uninfected saliva sample spiked with SARS-CoV-2 spike proteins.*

The signal obtained from the positive samples was normalised with respect to that of the corresponding antibody concentration in the PBS. **Figure 5.5 a** shows that a 50 ng mL^{-1} antibody concentration gives the highest average response from a viral load of 10^3 to 10^6 copies mL^{-1} compared to the other antibody concentrations. Hence, this was selected as the optimal concentration for the labelling process as it provides nearly optimal discrimination between positive samples and the antibody signal.

Next, we investigated the incubation time required for the labelling protocol. For this, two positive SARS-CoV-2 saliva samples, with viral loads 4×10^4 and 1.6×10^8 copies mL^{-1}, and one healthy donor sample spiked with recombinant spike proteins were tested at different incubation time points. The results plotted in **Figure 5.5 b** show a monotonic trend for all the samples in the first 20 minutes. The SARS-CoV-2 positive sample with 4×10^4 copies mL^{-1} reaches a plateau after 20 minutes and shows only a small increase in the signal after a prolonged incubation period. This effect may be related to the steady formation of antibody-antigen complexes. During the same time period, however, the signal in the sample with the highest viral load (1.6×10^8 copies mL^{-1})

decreases, indicating a Hook effect[32]. This implies that the antibody's binding sites may have been saturated due to the presence of an excess number of viral antigens. Thus, from this result the optimum labelling time was found to be around 20 minutes.

To further confirm that the increase in signal obtained with the FVR is mainly due to aggregation formed between the viral antigens and the antibody, we performed stimulated emission depletion (STED) imaging on a labelled SARS-CoV-2 positive saliva sample and also on an anti-SARS-CoV spike antibody only. For this, we used coverslips coated with poly-L-Lysine at 0.01% in MiliQ to improve the adhesion of the viral particles. 20 µL of labelled virus solution and the anti-SARS-CoV spike antibody were spotted on the coated coverslips and left to adhere for 15-20 minutes, after which they were imaged using LEICA TCS SP8 X. It can be observed from **Figure 5.6** that there is a 5-fold increase in the fluorescence signal of the labelled viral particles with respect to the background, whereas there is only a 1-fold increase in that of the antibodies.
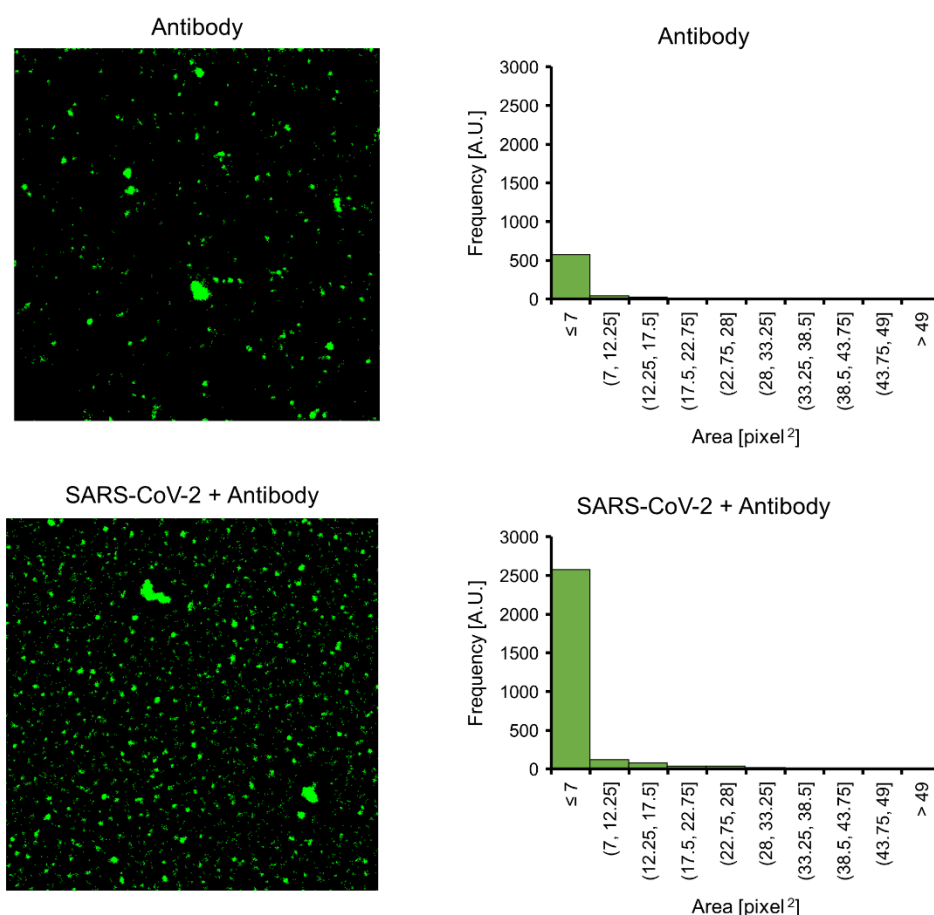
Figure 5.6: Super-resolution microscopy, STED images of anti-SARS-CoV spike antibodies only (top) and labelled SARS-CoV-2 positive saliva sample (bottom). *Corresponding plots of frequency vs. area show (bottom) a 5-fold increase in the fluorescence signal in the labelled SARS-CoV-2 positive saliva sample, and (top) only 1-fold increase in fluorescence signal for the anti-SARS-CoV spike antibodies. The analysis was performed with ImageJ 1.53c.*

5.5 Analytical Characterisation of the FVR

5.5.1 Limit of Blank (LoB) and Cut-off

First, we examined the signal from the filtered saliva samples and the Alexa 488-anti-SARS-CoV spike antibody in PBS (**Figure 5.7 a**). It can be seen that the signal obtained from the filtered saliva samples is lower than that from the Alexa 488-anti-SARS-CoV spike antibody ($p = 0.0001$), demonstrating that the noise from the remaining saliva content after filtration is negligible and has no effect on SARS-CoV-2 detection.

The Limit of Blank (LoB) of the FVR is calculated using the signal obtained from the Alexa 488-anti-SARS-CoV spike antibody signal. When samples containing no SARS-CoV-2 antigens are tested, the LoB represents the highest apparent SARS-CoV-2 concentration detected by the FVR. In our case, the LoB is defined as the average of the blank plus 1.645 times its standard deviation[33]. For the detection of SARS-CoV-2 positive samples, we set the cut-off value higher than the LoB, specifically, as the average of the blank (antibody signal) plus two times its standard deviation (**Figure 5.7 a**). Importantly, LoB and cut-off values must be established each time a new batch of the fluorescently labelled SARS-CoV-2 spike proteins is measured.

5.5.2 Standard Curve and LOD

In order to establish the relationship between the signal acquired from the FVR and the viral load in the tested samples, we built a standard curve by measuring triplicates of labelled SARS-CoV-2 viral antigens from saliva samples with increasing viral load concentrations. Three SARS-CoV-2 positive samples, with an initial virus load of 10^9 copies mL^{-1} (measured by RT-qPCR), were first diluted 100-fold in PBS. Further dilutions were prepared from each of these samples to yield concentrations ranging from 10^2 to 10^7 copies mL^{-1}.
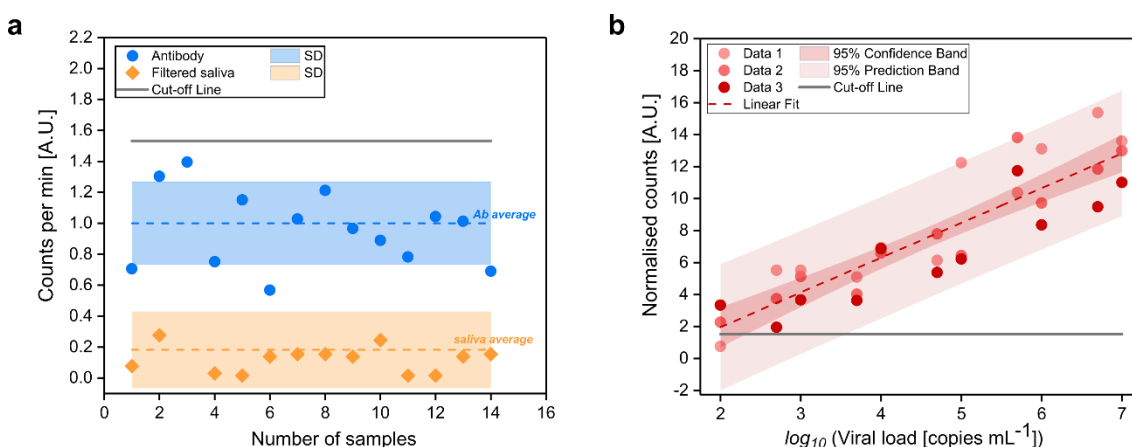
Figure 5.7: Determination of FVR cut-off value and standard curve for SARS-CoV-2 detection. (a) Counted events per minute of the Alexa 488-anti-SARS-CoV spike antibody (blue) and of filtered saliva samples (orange). Mean values (dotted lines) and standard deviations (shaded area) are indicated. Cut-off, according to the antibody signal (blue), was calculated as mean plus two times its standard deviation and is indicated as a solid line. (b) Standard curve of the SARS-CoV-2 viral particle concentration against normalised counts; $R^2 = 0.94$.

In **Figure 5.7 b,** the counted events of each measured sample are normalised to the antibody average signal to account for batch effects and user-to-user variability in sample preparation. The

standard curve shows strong linearity with $R^2 = 0.94$ and $p < 0.001$. The sensitivity calculated from the slope of the standard curve is 2.17 A.U. mL copies^{-1}, demonstrating the competence of the newly developed FVR in screening SARS-CoV-2 viral particles.

The limit of detection (LOD) of the FVR is calculated as the LOB plus 1.645 times the standard deviation of the sample with the lowest concentration of the SARS-CoV-2 viral particles tested[34]. According to the standard curve, the LOD of our FVR is found to be as low as 3,834 copies mL^{-1}, which is at least three orders of magnitude lower than commercially available rapid antigen tests[35]. The analytical range in which the FVR's response shows a monotonic trend is determined to be from 4×10^3 to 10^7 viral load mL^{-1}.

5.6 SARS-CoV-2 Detection Using FVR via Blind Test

To investigate the potential of our FVR as a novel platform for the fast detection of SARS-CoV-2 directly from saliva samples, we conducted a blind test using frozen saliva samples from SARS-CoV-2 infected (n=34) and uninfected (n=20) individuals (**Figure 5.8 a**). Without disclosing the results, these samples were first measured using RT-qPCR. Each sample was diluted 1:100 in PBS in order to reduce viscosity and to obtain a sufficient volume for filtering and measurement. The diluted samples were then filtered and labelled with an Alexa-488 anti-SARS-CoV spike antibody. As previously stated, the fluorescent antibody solution was initially measured as blank to determine the cut-off line, allowing the positive and negative SARS-CoV-2 samples to be distinguished in this way. The results plotted in **Figure 5.8 b** show that the FVR is able to reliably detect 31 out of 34 SARS-CoV-2 positive patient samples and 18 out of 20 SARS-CoV-2 negative patient samples from the samples previously analysed using RT-qPCR. According to the new European Commission's Medical Device Coordination Group's guidance document[36], this translates to a diagnostic sensitivity of 91.2 % and a diagnostic specificity of 90%. It's worth noting here that we selected a polyclonal antibody for our measurements due to its commercial availability. This antibody is elicited against the SARS-CoV spike glycoprotein, which shows cross-reactivity with the SARS-CoV-2 spike. In the future, we will focus on selecting antibodies that form complexes solely with the antigen of interest, such as monoclonal rather than polyclonal antibodies, to further improve the specificity of the assay.
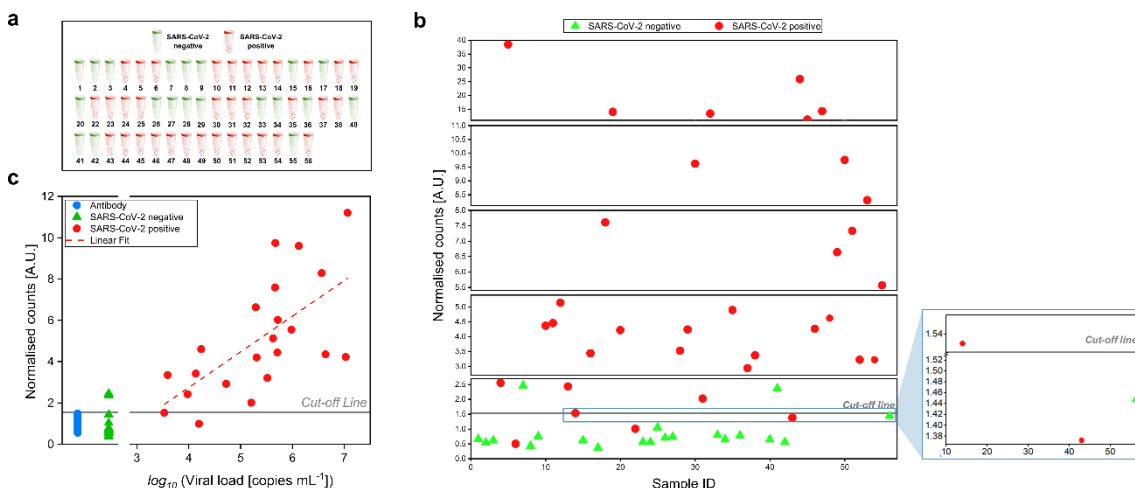
Figure 5.8: Detection of SARS-CoV-2 in saliva samples in a blind test. *(**a**) Sample identification by colour code (red: SARS-CoV-2 positive samples, n=34; green: SARS-CoV-2 uninfected samples, n=20) according to RT-qPCR. (**b**) Distribution of the saliva samples showing the normalised count obtained with the FVR.(**c**) Correlation between the normalised counts and log of viral load [copies mL^{-1}] by RT-qPCR of the blind test samples. The dashed line is a linear fit ($R^2 = 0.43$) to the data from the positive samples that fall within our analytical range, eliminating the outliers shown by a circle. Outliers are defined as data points with residual values larger than two times the standard deviation of the entire sample.*

In **Figure 5.8 c**, the signal from the FVR for the blind samples plotted against the log of viral load quantified by the RT-qPCR test shows a linear relationship ($R^2 = 0.43$) within the analytical range ($\sim 10^3 - 10^7$ copies mL^{-1}) of our FVR. However, this linearity is not maintained over the full viral load range of the samples tested. More specifically, the response is linear and monotonically increases for low concentrations. For concentrations higher than 10^7 copies mL^{-1}, the signal starts decreasing, almost linearly. A possible explanation for this behaviour is that the FVR response is dependent on the antigen concentration. Taking a closer look at the shape of the curve in **Figure 5.9**, we can see two distinct regions: (i) an antibody excess region and (ii) an antigen excess region. The antigens on the viral particles act as a bridge in the antibody excess region, coagulating many labelled antibodies, resulting in an increase in the signal compared to the blank solution, and, also, an increase in the signal when viral particle concentration rises. The signal begins to decrease in the antigen excess region, most likely because the bridging process is unfavourable and the signal from the labelled antibodies is masked by the presence of excess viral particles.
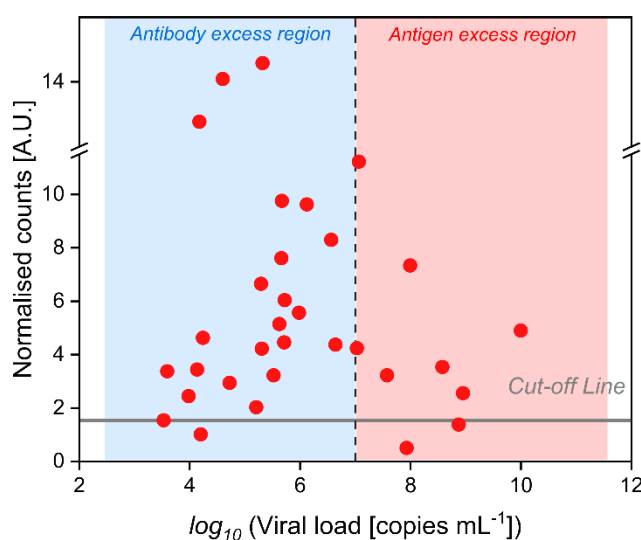
Figure 5.9: FVR normalised counts vs. RT-qPCR for the full viral load range measured in the blind test. The red dots represent the normalised counts from positive SARS-CoV-2 saliva samples measured with the FVR plotted against the log of the viral load determined by RT-qPCR. The grey line indicates the cut-off line. The shaded blue area is the antibody excess region and the shaded red area is the antigen excess region. A vertical black dashed line divides the antibody and antigen excess regions.

5.7 Quantification of SARS-CoV-2 in Saliva Samples

The viral load quantification is performed using the standard curve with equation y=2.1742x-2.3842. To account for the 1:100 dilution ratio used to prepare the blind test samples, the obtained viral load estimate is multiplied by 100. The statistical relationship between the viral load detected by RT-qPCR and our FVR, calculated using a Pearson's correlation test (**Figure 5.10 a**), shows a good correlation, with a coefficient r of 0.66 and a *p* value of 0.001 for the blind samples lying within our analytical measurement range (10^7 copies mL^{-1}). We are unable to accurately quantify the viral load in samples that fall above the FVR's analytical limit, but we are still able to establish whether the tested sample is positive or negative. In order to better understand the correlation between the FVR and the RT-qPCR, a Bland-Altman plot is shown in **Figure 5.10 b** to examine the agreement of the two procedures.
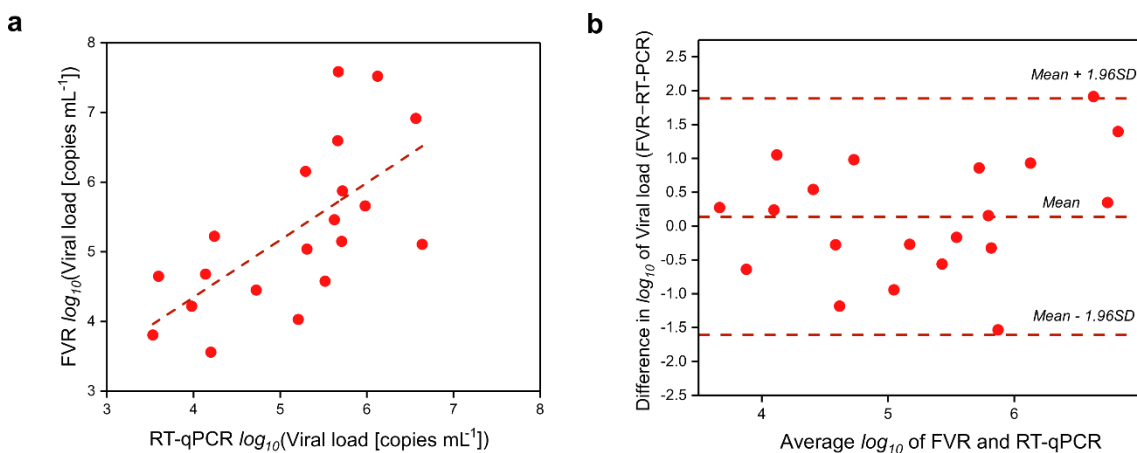
Figure 5.10: FVR vs RT-qPCR. (a) *Red dots represent the viral load quantification with the FVR reader versus the RT-qPCR viral load. Dashed red line is the linear regression of the quantification curve. The equation of the linear regression is y=0.8216x + 1.0612 (R^2=0.43); **(b)** Bland-Altman plot showing the difference in the viral load prediction between the FVR reader and the RT-qPCR. The normalised root-mean-square deviation NRMSD associated to the FVR predicted viral load is 9.2%.*

Considering that the differences in the measurements between the two methods have a normal distribution, we selected the limit of agreement as the mean of the differences in the measurements plus 1.96 times its standard deviation. The average of the differences is found to be 0.138, meaning that the FVR measures, on average, 0.138 times more than the log10 of the viral load with respect to RT-qPCR.

5.8 Conclusion

In this chapter, we have demonstrated the development of a novel flow-virometry reader (FVR) for fast, inexpensive and reliable SARS-CoV-2 diagnosis directly from saliva samples combining the short time to results of an antigen test and a high sensitivity, close to that of RT-PCR (**Table 5.1**). The FVR was designed using a small microfluidic chip combined with a laser-pumped optical head to detect the presence of viruses directly from saliva samples tagged with fluorescent antibodies. For sample preparation, we developed a simple protocol based on saliva collection via spit, purification through filtration and labelling of the saliva samples with an Alexa-488 anti-SARS-CoV spike antibody. The main advantage of such a protocol using saliva is that it does not require a dedicated lab, causes less discomfort to patients and can even be self-administered, unlike the nasal or nasopharyngeal swabs required for RT-PCR. Furthermore, the saliva was collected by directly spitting into a sterile tube without the use of the Salivette® test kit as is used in other detection techniques and without prior restrictions on the intake of food or drinks. Also, there is no need to purify the sample after labelling because the number of detected events in a positive sample is substantially higher than the antibody signal. Since efficient antigen-antibody binding is critical for detecting SARS-CoV-2 with the FVR, the labelling technique was first optimised by determining the optimal antibody concentration and incubation time required to prepare the samples. Next, a standard curve was created using clinical saliva samples with varying viral loads from SARS-CoV-2 positive patients confirmed by RT-qPCR. A low LOD of 3,834 copies mL^{-1} was achieved and the FVR's analytical range was established to be from ~ 10^3 to 10^7 viral copies mL^{-1}. In a blind test with 54 clinical samples, the reader was found to have a high sensitivity of 91.2 % and a high specificity of 90 %. Furthermore, we were able to predict the concentrations of the unknown samples (copies mL^{-1}) based on our standard curve and compare them to the RT-qPCR results. After discarding the samples that returned a normalised count over the upper limit of our calibration curve, we were able to establish the viral load mL^{-1} of the measured samples within the detection range of the FVR, allowing a normalised root-mean-square deviation (NRMSD) of 9.2 %.

When comparing the above-mentioned FVR specifications with respect to other detection techniques that received FDA Emergency Use Authorization, it can be seen from Table 5.1 that in terms of sensitivity, FVR is similar to RT-PCR and RAT while slightly better than RT-LAMP depending on the manufacturer. The specificity falls within the lower end of RT-PCR while RT-LAMP and RAT report a higher specificity compared to FVR. Therefore, future work will focus on improving the specificity of the assay. In terms of LOD, FVR is comparable to RT-PCR's 10^3-10^5 copies mL^{-1} and RT-LAMP's 10^3-7.5×10^5 copies mL^{-1}, respectively, while at least one to three orders of magnitude lower than RAT's 10^4-10^7 copies mL^{-1} (approximate conversion).

Table 5.1: A comparison of commercial tests for detection of SARS-CoV-2[10].

Type of Test	Detection Method	Specimen	Time to Results	Sensitivity*	Specificity*	Limit of Detection (LOD)[†]	Assay Complexity**	Portability
RT-PCR	Molecular assay for qualitative/ quantitative detection of specific gene sequences of SARS-CoV-2	Upper and lower respiratory specimens, sputum, saliva	2-4 h	87.5% – 100%	90% - 100%	$1 - 100$ copies/ µL	Middle-to-High	Yes
RT-LAMP	Molecular assay for qualitative detection of specific gene sequences of SARS-CoV-2	Anterior nasal/ nasopharyngeal swab, sputum, saliva	15-60 min	93% – 100%	98% - 100%	$1 - 75$ copies/µL	Middle	Yes
Rapid Antigen Tests (RAT)	Lateral flow immunoassay for qualitative detection of SARS-CoV-2 nucleocapsid protein	Anterior nasal/ nasopharyngeal swab	15 – 30 min	80% - 97%	96.6% - 100%	$10^1 - 10^4$ TCID$_{50}$/mL	Low	Yes
FVR	**Flow virometry for qualitative/ quantitative detection of SARS-CoV-2 spike protein**	**Saliva**	**25 – 30 min**	**91%**	**90%**	**3,834 copies/mL**	**Low-to-Middle**	**Yes**

* Depending on manufacturer/ sample type/ collection

** Low= does not require the use of pipette and it needs little operational steps.

Middle = it requires the use of pipette and might require the use of additional specialized equipment (other than the detection system).

High = it requires the use of trained personnel, pipette and additional specialized equipment (other than the detection system).

[†] Various assays report LODs in different units. Here, we use the unit specified by the manufacturer without converting to copies/mL to be more accurate. Units such as Median Tissue Culture Infectious Dose, TCID$_{50}$, may be related to viral copies in different ways depending on the viral preparation.

Based on the above findings, the FVR has great potential to bridge the gap between RAT and RT-PCR tests, allowing for a fast, portable, user-friendly, and novel point-of-care device that can run up to 2000 tests per day. Its throughput could be substantially improved by multiplexing numerous samples within the same device and executing them all at the same time. In terms of cost, the current FVR has a bill of materials of approximately 10,000 € since it is built with off-the-shelf components. The overall cost of the system will be greatly reduced when transitioning from the current prototype to large-scale manufacturing. Saliva collection in that case could be performed

using a self-collection kit with a tube pre-filled with reagents and a filter to clarify the saliva. It is conceivable that the operational cost of a single test will be as low as 2 €. Furthermore, the technology can readily be extended to identify other possible variants of SARS-CoV-2 or other particular viruses, such as seasonal coronaviruses or the influenza virus, with correct antibody selection. To this end, future research should focus on antibody cross-reactivity investigations with other viruses and antigens. Additionally, efforts will be made to establish an integrated FVR for simultaneous antigen and antibody analysis of SARS-CoV-2. This combined testing device will be able to identify not only an active infection in the body (antigen detection), but also information on the immunological response to the infection or after vaccination (antibody detection). This will allow for a rapid evaluation of vaccine efficacy and the requirement for subsequent vaccination boosters.

5.9 References

1. Rybicki, E. The classification of organisms at the edge of life, or problems with virus systematics. *S. Afr. J. Sci.* **86**, 182–186 (1990).

2. S, M.-C., MK, J. & AM, M. COVID-19 and globalization. *One Heal. (Amsterdam, Netherlands)* **9**, (2020).

3. McKee, M. & Stuckler, D. If the world fails to protect the economy, COVID-19 will damage health not just now but also in the future. *Nat. Med. 2020 265* **26**, 640–642 (2020).

4. Wei, S. *et al.* Field-deployable, rapid diagnostic testing of saliva for SARS-CoV-2. *Sci. Reports 2021 111* **11**, 1–9 (2021).

5. Xun, G., Lane, S. T., Petrov, V. A., Pepa, B. E. & Zhao, H. A rapid, accurate, scalable, and portable testing system for COVID-19 diagnosis. *Nat. Commun. 2021 121* **12**, 1–9 (2021).

6. Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565–574 (2020).

7. Park, M., Won, J., Choi, B. Y. & Lee, C. J. Optimization of primer sets and detection protocols for SARS-CoV-2 of coronavirus disease 2019 (COVID-19) using PCR and real-time PCR. *Exp. Mol. Med.* **52**, 963–977 (2020).

8. Bohn, M. K. *et al.* Molecular, serological, and biochemical diagnosis and monitoring of COVID-19: IFCC taskforce evaluation of the latest evidence. *Clin. Chem. Lab. Med.* **58**, 1037–1052 (2020).

9. Ji, T. *et al.* Detection of COVID-19: A review of the current literature and future perspectives. *Biosens. Bioelectron.* **166**, (2020).

10. Antibody, Antigen and Molecular Tests for COVID-19. *JOHNS HOPKINS CENTER FOR HEALTH SECURITY* (2021).

11. WE, H. *et al.* RT-LAMP for rapid diagnosis of coronavirus SARS-CoV-2. *Microb. Biotechnol.* **13**, 950–961 (2020).

12. Li, E., Larson, A., Kothari, A. & Prakash, M. Handyfuge-LAMP: low-cost and electricity-free centrifugation for isothermal SARS-CoV-2 detection in saliva. *medRxiv* 2020.06.30.20143255 (2020) doi:10.1101/2020.06.30.20143255.

13. Pilarowski, G. *et al.* Field performance and public health response using the BinaxNOWTM Rapid SARS-CoV-2 antigen detection assay during community-based testing. *Clin Infect Dis* 1–27 (2020).

14. Krüger, L. J. *et al.* Evaluation of the accuracy, ease of use and limit of detection of novel, rapid, antigen-detecting point-of-care diagnostics for SARS-CoV-2. *medRxiv* (2020) doi:10.1101/2020.10.01.20203836.

15. Kevadiya, B. D. *et al.* Diagnostics for SARS-CoV-2 infections. *Nat. Mater.* **20**, 593–605 (2021).

16. Ji, T., Liu, Z., Wang, G., Guo, X. & Akbar, S. Detection of COVID-19: A review of the current literature and future perspectives. *Biosens Bioelectron* **166**, 112455 (2020).

17. Vandenberg, O., Martiny, D., Rochas, O., van Belkum, A. & Kozlakidis, Z. Considerations for diagnostic COVID-19 tests. *Nat. Rev. Microbiol.* (2020) doi:10.1038/s41579-020-00461-z.

18. Sin, M. N. et al. Microbiological Guidelines for Food. *Cent. Food Saf.* **2014**, 1–38 (2014).

19. Alpdagtas, S. *et al.* Evaluation of current diagnostic methods for COVID-19. *APL Bioeng.* **4**, 41506 (2020).

20. Chen, Q. *et al.* Diagnostic technologies for COVID-19: a review. *RSC Adv.* **10**, 35257–35264 (2020).

21. B, G. *et al.* Review of analytical performance of COVID-19 detection methods. *Anal. Bioanal. Chem.* **413**, 35–48 (2021).

22. Asghari, A. *et al.* Fast, accurate, point-of-care COVID-19 pandemic diagnosis enabled through advanced lab-on-chip optical biosensors: Opportunities and challenges. *Appl. Phys. Rev.* **8**, 031313 (2021).

23. Ahmad Bhutta, Z. *et al.* Nanophotonic biosensors for point-of-care COVID-19 diagnostics and coronavirus surveillance. *J. Phys. Photonics* **3**, 011002 (2021).

24. Goswami, N. *et al.* Label-free SARS-CoV-2 detection and classification using phase imaging with computational specificity. *Light Sci. Appl. 2021 101* **10**, 1–12 (2021).

25. Wyllie, A. L. *et al.* Saliva is more sensitive for SARS-CoV-2 detection in COVID-19 patients than nasopharyngeal swabs. *medRxiv* (2020) doi:10.1101/2020.04.16.20067835.

26. Li, Y. *et al.* Saliva is a non-negligible factor in the spread of COVID-19. *Mol. Oral Microbiol.* **35**, 141–145 (2020).

27. Sapkota, D., Thapa, S. B., Hasséus, B. & Jensen, J. L. Saliva testing for COVID-19? *Br. Dent. J.* **228**, 658–659 (2020).

28. McKinnon, K. M. Flow Cytometry: An Overview. *Curr. Protoc. Immunol.* **120**, 5.1.1 (2018).

29. Zamora, J. L. R. & Aguilar, H. C. Flow virometry as a tool to study viruses. *Methods* vols 134–135 87–97 (2018).

30. Reverberi, R. & Reverberi, L. Factors affecting the antigen-antibody reaction. *Blood Transfus.* **5**, 227–240 (2007).

31. Batéjat, C., Grassin, Q., Manuguerra, J.-C. & Leclercq, I. Heat inactivation of the severe acute respiratory syndrome coronavirus 2. *J. Biosaf. Biosecurity* **3**, 1–3 (2021).

32. Ward, G., Simpson, A., Boscato, L. & Hickman, P. E. The investigation of interferences in immunoassay. *Clin. Biochem.* **50**, 1306–1311 (2017).

33. Armbruster, D. A. & Pry, T. Limit of blank, limit of detection and limit of quantitation. *Clin. Biochem. Rev.* **29 Suppl 1**, S49-52 (2008).

34.   Shrivastava, A. & Gupta, V. Methods for the determination of limit of detection and limit of quantitation of the analytical methods. *Chronicles Young Sci.* **2**, 21 (2011).

35.   Corman, V. M. *et al.* Comparison of seven commercial SARS-CoV-2 rapid point-of-care antigen tests : a single-centre laboratory evaluation study. *The Lancet Microbe* **2**, 311–319 (2021).

36.   Medical Device Coordination Group. *MDCG 2021-21 Guidance on performance evaluation of SARS-CoV-2 in vitro diagnostic medical devices.* (2021).

CHAPTER 6

Summary and Outlook

This thesis was aimed at contributing to the field of particle detection and characterisation by developing robust, portable and inexpensive photonic sensors with adequate performance for a variety of industrial and biological applications.

In Chapter 2, we discussed the necessity for particle characterisation, specifically the determination of particle size and counts. We presented a general overview of the most widely used and commercially available particle characterisation techniques and instruments, together with their characteristics, advantages, and limitations.

Optical techniques based on light scattering have been shown to outperform traditional methods, particularly in determining the particle size distribution of samples, due to their high accuracy, short measurement time, and non-invasive detection principle. This formed the basis for the work presented in Chapter 3, in which we introduced the concept of a new design of a simple, compact, and reliable particle size analyser (PSA). Our device includes an innovative element, which is a patented angular spatial filter (ASF), together with a single CMOS image sensor, and off-the-shelf electronics. By combining the hardware with a customised machine learning (ML) algorithm, we demonstrated that the proposed system can measure particles ranging from approximately 10 µm to 100 µm in size with a mean absolute percentage error (MAPE) of only 0.72% for spherical particles. This performance is comparable to that specified by commercial laser diffraction (LD) based PSAs. We also presented the preliminary results from a flow-through experiment, which further showed the PSA's potential as an integrable online or at-line platform for quality control monitoring in industries such as pharmaceuticals and food, among others.

The work detailed in Chapter 4 was aimed at improving the performance of the ASF by reducing residual spurious effects, such as internal reflections from its inner sidewalls, and incorporating this into an optimised PSA design capable of measuring particles smaller than 10 µm. We accomplished this by removing the forward unscattered light from the sample suspension, leading to an increased system sensitivity and capability to distinguish small particles, particle refractive index contrast and shape. When combined with a proper classification ML model, the PSA can distinguish between *E. coli*, *Enterococcus sp.* and silica particles in water with an accuracy of 89% within the PSA's optimal operating concentration range. Furthermore, the findings obtained from testing deionized water with different bacterial concentrations using a regression ML model demonstrated the device's potential in predicting concentrations of *E. coli* and *Enterococcus sp.* with a mean absolute error (MAE) of only 0.36 and 0.16 within the limit of detection (LOD) of the system.

For the specific detection of viruses, which are smaller than 1 µm, direct light scattering measurements alone are not sufficient. We addressed this in Chapter 5, where we described the concept behind the development of a portable and cost-effective flow-virometry reader (FVR) that takes advantage of the fluorescence detection mechanism of a traditional flow-cytometer

while overcoming its limitations in POC applications. To demonstrate a use case, we customised the FVR for the rapid detection of a SARS-CoV-2 complex sample, such as saliva, without the need for any laboratory equipment, e.g., a centrifuge. The measurement time was less than 30 minutes, and we obtained an LOD of 3,834 copies mL^{-1}, which is at least three orders of magnitude lower than the commercially available antigen tests on the market. The analytical measuring range was from ~ 10^3 to 10^7 viral copies mL^{-1}. We validated the system in a blind test with 54 clinical samples and achieved a high sensitivity of 91.2% and a high specificity of 90%. Furthermore, when compared with RT-qPCR results, we demonstrated that we were able to quantify the viral load in the unknown saliva samples within the detection range of the FVR.

The preliminary findings from the compact photonic sensors developed in this thesis show that these devices are powerful tools for particle characterisation with great potential for performing large-scale and high throughput measurements. This thanks to their low cost, portability, reduced testing time and simple measurement techniques that do not require specialised equipment or trained personnel. As a result of these benefits, the implementation of such devices will be particularly useful in medium to low-income countries, for example in rural areas or hard-to-reach locations where resources, budget, and skilled staff are limited.

Outlook

The devices described in this thesis were validated in controlled laboratory environments, and further research and development are required before the technologies can be fully exploited for practical applications in real environments. Future work will focus on several aspects which are outlined below.

- The low prediction error obtained with the PSA for large particle measurement is for median diameter D50 only. In industrial applications, particle suspensions are often polydisperse in nature; therefore, defining a single average diameter does not provide a complete size description of the measured sample. Although the developed ML models were used to estimate D10 and D90 values to assess the particle size distribution, the prediction errors obtained from the initial results need to be improved further. Efforts will be focused on lowering the prediction error by including more data with diverse size distributions in the ML analysis.

- When using the modified PSA for biological applications, it would be interesting to investigate the potential of the system in distinguishing other types of pathogenic bacteria, in addition to *E. coli* and *Enterococci*. Another route to explore is labelling *E. coli* with antibody-coated microbeads or alternatively magnetic beads for specific detection. In the former case, this could help further improve the limit of detection by enhancing light scattering from bacteria at low concentrations.

- The FVR specificity in detecting SARS-CoV-2 may be improved by carefully selecting antibodies that only form complexes with the target antigen. In addition, with additional limited hardware changes, a similar approach may lead to the employment of the FVR for the detection of other viruses. The LOD could also be improved by increasing the sample size or adjusting the antibody-antigen ratio, for example. Moreover, further developments will be made to the microfluidic components to potentially allow simultaneous measurements of multiple samples. With these modifications, the system may reach a configuration for clinical trials in various settings, such as hospital emergency rooms, care homes, etc.

While the proposed devices have shown potential for applications, further work is also needed to elucidate which configurations are more likely to have a commercial impact and generate a product. To this end, engagement with end-users and customers is needed, and the devices will have to be customised for their specific applications. This will also require engineering work to make reliable working prototypes outside of the laboratory.

Bibliography

1. Shekunov, B. Y., Chattopadhyay, P., Tong, H. H. Y. & Chow, A. H. L. Particle size analysis in pharmaceutics: Principles, methods and applications. *Pharm. Res.* **24**, 203–227 (2007).

2. Xu, R. Light scattering: A review of particle characterization applications. *Particuology* vol. 18 11–21 (2015).

3. Castillo-Henríquez, L. *et al.* Biosensors for the Detection of Bacterial and Viral Clinical Pathogens. *Sensors* **20**, 6926 (2020).

4. Parrish, K. & Fahrenfeld, N. L. Microplastic biofilm in fresh- and wastewater as a function of microparticle type and size class. *Environ. Sci. Water Res. Technol.* **5**, 495–505 (2019).

5. Harrison, R. M. Key pollutants—airborne particles. *Sci. Total Environ.* **334–335**, 3–8 (2004).

6. Narayanaswamy, R. & Wolfbeis, O. S. *Optical Sensors. Medical Devices and Systems* vol. 1 (Springer Berlin Heidelberg, 2004).

7. Blott, S. J., Croft, D. J., Pye, K., Saye, S. E. & Wilson, H. E. Particle size analysis by laser diffraction. *Geol. Soc. London, Spec. Publ.* **232**, 63–73 (2004).

8. Barber, T. A. *Control of Particulate Matter Contamination in Healthcare Manufacturing. Control of Particulate Matter Contamination in Healthcare Manufacturing* (CRC Press, 1999). doi:10.1201/9780429246692.

9. Gillespie, S. *et al.* Assessing microbiological water quality in drinking water distribution systems with disinfectant residual using flow cytometry. *Water Res.* **65**, 224–234 (2014).

10. Yang, L. & Yamamoto, T. Quantification of Virus Particles Using Nanopore-Based Resistive-Pulse Sensing Techniques. *Front. Microbiol.* **7**, 1500 (2016).

11. Merkus, H. G. Particle Size, Size Distributions and Shape. in *Particle Size Measurements* 13–42 (Springer Netherlands, 2009). doi:10.1007/978-1-4020-9016-5_2.

12. Jennings, B. R. & Parslow, K. Particle size measurement: the equivalent spherical diameter. *Proc. R. Soc. London. A. Math. Phys. Sci.* **419**, 137–149 (1988).

13. Staniforth, J. N. & Taylor, K. M. G. Particle size analysis | Basicmedical Key. https://basicmedicalkey.com/particle-size-analysis/ (2016).

14. Allen, T. Particle size analysis by sieving. in *Powder Sampling and Particle Size Determination* 208–250 (Elsevier, 2003). doi:10.1016/B978-044451564-3/50006-1.

15. Sieve Analysis - Particle Technology Labs. https://www.particletechlabs.com/analytical-testing/particle-size-distribution-analyses/sieve-analysis (2021).

16. Sifting Through Sieve Analysis | Innopharma Technology. https://www.innopharmatechnology.com/news/sifting-through-sieve-analysis.

17. Allen, T. *Particle Size Measurement*. (Springer US, 1981). doi:10.1007/978-1-4899-3063-7.

18. Calvert, J. G. Glossary of atmospheric chemistry terms (Recommendations 1990). *Pure Appl. Chem.* **62**, 2167–2219 (1990).

19. Allen, T. A Review of Sedimentation Methods of Particle Size. in *Particle Size Analysis* 454 (Royal Society of Chemistry, 2007).

20. Figueiredo, M. M. Electrozone Sensing in Particle Size Analysis. in *Encyclopedia of Analytical Chemistry* (John Wiley & Sons, Ltd, 2006). doi:10.1002/9780470027318.a1504.

21. Horák, D., Peška, J., Švec, F. & Štamberg, J. The influence of porosity of discrete particles upon their apparent dimensions as measured by the Coulter principle. *Powder Technol.* **31**, 263–267 (1982).

22. Allen, T. *Powder Sampling and Particle Size Determination. Powder Sampling and Particle Size Determination* (Elsevier, 2003). doi:10.1016/B978-0-444-51564-3.X5000-1.

23. Bohren, C. F. & Huffman, D. R. *Absorbing and scattering of light by small particles. Wiley Interscience* (1983). doi:10.1002/9783527618156.

24. Mie - scatterlib. http://scatterlib.wikidot.com/mie.

25. Hulst, H. C. van de. Light scattering by small particles. By H. C. van de Hulst. New York (John Wiley and Sons), London (Chapman and Hall), 1957. Pp. xiii, 470; 103 Figs.; 46 Tables. 96s. *Q. J. R. Meteorol. Soc.* **84**, 198–199 (1958).

26. Black, D. L., McQuay, M. Q. & Bonin, M. P. Laser-based techniques for particle-size measurement: A review of sizing methods and their industrial applications. *Prog. Energy Combust. Sci.* **22**, 267–306 (1996).

27. Ma, Z., Merkus, H. G., De Smet, J. G. A. E., Heffels, C. & Scarlett, B. New developments in particle characterization by laser diffraction: size and shape. *Powder Technol.* **111**, 66–78 (2000).

28. 13320. Particle size analysis-laser diffraction methods. *ISO Stand. Auth.* (2009).

29. Laser Diffraction Technology for Routine Particle Size Analysis. 1–6 https://www.azom.com/article.aspx?ArticleID=11363 (2015).

30. Ye, M. *et al.* Inversion of particle-size distribution from angular light-scattering data with genetic algorithms. *Appl. Opt.* **38**, 2677 (1999).

31. Twomey, S. On the Numerical Solution of Fredholm Integral Equations of the First Kind by the Inversion of the Linear System Produced by Quadrature. *J. ACM* **10**, 97–101 (1963).

32. Vargas-Ubera, J., Aguilar, J. F. & Gale, D. M. Reconstruction of particle-size distributions from light-scattering patterns using three inversion methods. *Appl. Opt. Vol. 46, Issue 1, pp. 124-132* **46**, 124–132 (2007).

33. Liu, J. Essential parameters in particle sizing by integral transform inversions. *Appl. Opt.* **36**, 5535 (1997).

34. Jones, A. R. Light scattering for particle characterization. *Prog. Energy Combust. Sci.* **25**, 1–53 (1999).

35. Jonasz, M. Size, shape, composition, and structure of microparticles from light scattering. in *Principles, Methods and Application of Particle Size Analysis* 143–162 (Cambridge University Press, 1991). doi:10.1017/CBO9780511626142.014.

36. Stetefeld, J., McKenna, S. A. & Patel, T. R. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys. Rev.* **8**, 409–427 (2016).

37. The principles of dynamic light scattering :: Anton Paar Wiki. https://wiki.anton-paar.com/at-de/das-prinzip-der-dynamischen-lichtstreuung/.

38. LS Instruments | Introduction. https://lsinstruments.ch/en/theory/dynamic-light-scattering-dls/introduction.

39. Ltd., M. I. *Achieving high sensitivity at different scattering angles with different optical configurations.* (2014).

40. Kim, A., Ng, W. B., Bernt, W. & Cho, N. J. Validation of Size Estimation of Nanoparticle Tracking Analysis on Polydisperse Macromolecule Assembly. *Sci. Rep.* **9**, 1–14 (2019).

41. Kim, A., Bernt, W. & Cho, N.-J. Improved Size Determination by Nanoparticle Tracking Analysis: Influence of Recognition Radius. *Anal. Chem.* (2019).

42. Picot, J., Guerin, C. L., Le Van Kim, C. & Boulanger, C. M. Flow cytometry: retrospective, fundamentals and recent instrumentation. *Cytotechnology* **64**, 109–130 (2012).

43. McKinnon, K. M. Flow Cytometry: An Overview. *Curr. Protoc. Immunol.* **120**, 5.1.1 (2018).

44. Flow Cytometry Guide - Creative Diagnostics. https://www.creative-diagnostics.com/flow-cytometry-guide.htm.

45. Yu, W. & Hancock, B. C. Evaluation of dynamic image analysis for characterizing pharmaceutical excipient particles. *Int. J. Pharm.* **361**, 150–157 (2008).

46. Whiting, J. G., Tondare, V. N., Scott, J. H. J., Phan, T. Q. & Donmez, M. A. Uncertainty of particle size measurements using dynamic image analysis. *CIRP Ann.* **68**, 531–534 (2019).

47. Valsangkar, A. J. Principles, methods and applications of particle size analysis. *Can. Geotech. J.* **29**, 1006–1006 (2008).

48. Servais, C., Jones, R. & Roberts, I. The influence of particle size distribution on the processing of food. *J. Food Eng.* **51**, 201–208 (2002).

49. Global Particle Size Analysis Market by Technology (Laser Diffraction, DLS, Imaging, Coulter Principle, Sieving, Nanoparticle Tracking), Dispersion (Wet, Dry, Spray), End-user (Pharma-biotech, Cosmeceutical, Chemicals, Food, Academia), and Region - Foreca. https://www.researchandmarkets.com/reports/5116160/global-particle-size-analysis-market-by?utm_source=GNOM&utm_medium=PressRelease&utm_code=qjrsvx&utm_campaign=1629282+-+Global+Particle+Size+Analysis+Market+(2021+to+2026)+-+Increasing+Number+of+Conference.

50. Vargas-ubera, J., Aguilar, J. F. & Gale, D. M. Light-Scattering Patterns Using Three Inversion Methods. (2007).

51. Ye, Z., Jiang, X. & Wang, Z. Measurements of particle size distribution based on Mie scattering theory and Markov chain inversion algorithm. *J. Softw.* **7**, 2309–2316 (2012).

52. Mishchenko, D. Travis, L. & Lacis, A. *Multiple Scattering of Light by Particles: Radiative Transfer and Coherent Backscattering*. (Cambridge University Press, 2006).

53. Lenke, R. & Maret, G. *Multiple scattering of light: Coherent backscattering and radiative transfer*. (Cambridge University Press, 2000).

54. Gomi, H. Multiple scattering correction in the measurement of particle size and number density by the diffraction method. *Appl. Opt.* **25**, 3552 (2009).

55. Quirantes, A., Arroyo, F. & Quirantes-Ros, J. Multiple light scattering by spherical particle systems and its dependence on concentration: A T-matrix study. *J. Colloid Interface Sci.* **240**, 78–82 (2001).

56. Wei, Y., Shen, J. & Yu, H. Numerical calculation of multiple scattering with the layer model. *Particuology* **7**, 76–82 (2009).

57. Harvill, T. L. & Holve, D. J. Method for measuring particle size in the presence of multiple scattering. (1997).

58. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).

59. Guardani, R., Nascimento, C. A. O. & Onimaru, R. S. Use of neural networks in the analysis of particle size distribution by laser diffraction: Tests with different particle systems. *Powder Technol.* **126**, 42–50 (2002).

60. C.A.O. Nascimento a. Use of neural networks in the analysis of particle size distributions by laser diffraction. *Powder Technol.* **90**, 89–94 (1997).

61. Breiman, L. Random Forests. *Mach. Learn. 2001 451* **45**, 5–32 (2001).

62. Pruneri, V., Martïnez Cordero, P. A. & Jofre Cruanyes, M. Apparatus for measuring light scattering. (2018).

63. Bohren, C. F. & Huffman, D. R. *Absorption and scattering of light by small particles.* (John Wiley & Sons, 2008).

64. Barton, G., Van Eijkelenborg, M. A., Henry, G., Large, M. C. J. & Zagari, J. Fabrication of microstructured polymer optical fibres. *Opt. Fiber Technol.* (2004) doi:10.1016/j.yofte.2004.05.003.

65. van Eijkelenborg, M. *et al.* Microstructured polymer optical fibre. *Opt. Express* **9**, 319 (2001).

66. Hastie, T., Tibshirani, R. & Friedman, J. Random Forests. in 587–604 (Springer, New York, NY, 2009). doi:10.1007/978-0-387-84858-7_15.

67. Hastie, T., Tibshirani, R., Friedman, J. & Franklin, J. The elements of statistical learning: data mining, inference and prediction. *Math. Intell.* **27**, 83–85 (2005).

68. Rajapaksha, P. *et al.* A review of methods for the detection of pathogenic microorganisms. *Analyst* **144**, 396–411 (2019).

69. Li, J., Zhu, Y., Wu, X. & Hoffmann, M. R. Rapid Detection Methods for Bacterial Pathogens in Ambient Waters at the Point of Sample Collection: A Brief Review. *Clin. Infect. Dis.* **71**, S84–S90 (2020).

70. Lazcka, O., Campo, F. J. Del & Muñoz, F. X. Pathogen detection: A perspective of traditional methods and biosensors. *Biosens. Bioelectron.* **22**, 1205–1217 (2007).

71. Chatterjee, A. & Abraham, J. Chapter 2 - Microbial Contamination, Prevention, and Early Detection in Food Industry. in *Microbial Contamination and Food Degradation* (eds. Holban, A. M. & Grumezescu, A. M.) 21–47 (Academic Press, 2018).

72. De Boer, E. & Beumer, R. R. Methodology for detection and typing of foodborne microorganisms. *Int. J. Food Microbiol.* **50**, 119–130 (1999).

73. Yasir, M. Analysis of Microbial Communities and Pathogen Detection in Domestic Sewage Using Metagenomic Sequencing. *Divers. 2021, Vol. 13, Page 6* **13**, 6 (2020).

74. Ajonina, C., Buzie, C., Rubiandini, R. H. & Otterpohl, R. Microbial pathogens in Wastewater Treatment Plants (WWTP) in Hamburg. *J. Toxicol. Environ. Heal. - Part A Curr. Issues* **78**, 381–387 (2015).

75. Noble, R. T. & Weisberg, S. B. A review of technologies for rapid detection of bacteria in recreational waters.

76. Law, J. W.-F., Mutalib, N.-S. A., Chan, K.-G. & Lee, L.-H. Rapid methods for the detection of foodborne bacterial pathogens: principles, applications, advantages and limitations. *Front. Microbiol.* **5**, (2014).

77. Valones, M. A. A. *et al.* Principles and applications of polymerase chain reaction in medical diagnostic fields: a review. *Brazilian J. Microbiol.* **40**, 1–11 (2009).

78. McCarthy, J. Immunological techniques: ELISA. in *Detecting Pathogens in Food* vol. 84 241–258 (Elsevier, 2003).

79. Hussain, M. *et al.* Rapid Identification of Pathogens based on MIE Light Scattering and Machine Learning Approach. *Med. Meas. Appl. MeMeA 2019 - Symp. Proc.* (2019) doi:10.1109/MEMEA.2019.8802228.

80. Liang, P. S., San Park, T. & Yoon, J.-Y. Light scattering-based detection of food pathogens. in *Light Scattering Technology for Food Property, Quality and Safety Assessment* 429–444 (CRC Press, 2017).

81. Bhunia, A. K., Banada, P., Banerjee, P., Valadez, A. & Hirleman, E. D. Light scattering, fiber optic- and cell-based sensors for sensitive detection of foodborne pathogens. *J. Rapid Methods Autom. Microbiol.* **15**, 121–145 (2007).

82. Sang, S. *et al.* Progress of new label-free techniques for biosensors: a review. *Crit. Rev. Biotechnol.* **36**, 1–17 (2015).

83. O, L., FJ, D. C. & FX, M. Pathogen detection: a perspective of traditional methods and biosensors. *Biosens. Bioelectron.* **22**, 1205–1217 (2007).

84. Cooper, M. A. Label-Free Biosensors: Techniques and Applications. *Label-Free Biosens. Tech. Appl.* 1–304 (2009).

85. Wieliczko, A., Podbielska, H. & Buzalewicz, I. Influence of various growth conditions on Fresnel diffraction patterns of bacteria colonies examined in the optical system with converging spherical wave illumination. *Opt. Express, Vol. 19, Issue 22, pp. 21768-21785* **19**, 21768–21785 (2011).

86. Abdelhaseib, M. U., Singh, A. K. & Bhunia, A. K. Simultaneous detection of Salmonella enterica, Escherichia coli and Listeria monocytogenes in food using a light scattering sensor. *J. Appl. Microbiol.* **126**, 1496–1507 (2019).

87. Haavig, D. L. *et al.* Cellular Light Scattering for the Identification of Bacteria and Its Application to the Identification of Staphylococcus. *J. AOAC Int.* **100**, 1836–1847 (2017).

88. Molleda, P., Blanco, I., Ansola, G. & de Luis, E. Removal of wastewater pathogen indicators in a constructed wetland in Leon, Spain. *Ecol. Eng.* **33**, 252–257 (2008).

89. García-Solache, M. & Rice, L. B. The enterococcus: A model of adaptability to its environment. *Clinical Microbiology Reviews* vol. 32 (2019).

90. Rompré, A., Servais, P., Baudart, J., De-Roubin, M. R. & Laurent, P. Detection and enumeration of coliforms in drinking water: Current methods and emerging approaches. *J. Microbiol. Methods* **49**, 31–54 (2002).

91. Subgingival Microbes. in *Atlas of Oral Microbiology* 67–93 (Elsevier, 2015). doi:10.1016/b978-0-12-802234-4.00004-5.

92. Gotoh, K., Masuda, H. & Higashitani, K. Powder technology handbook. *Journal of Thermal Spray Technology* vol. 10 10 (2001).

93. Smith, Z. J., Chu, K. & Wachsmann-Hogiu, S. Nanometer-Scale Sizing Accuracy of Particle Suspensions on an Unmodified Cell Phone Using Elastic Light Scattering. *PLoS One* **7**, e46030 (2012).

94. Jones, A. R. Light scattering for particle characterization. *Prog. Energy Combust. Sci.* **25**, 1–53 (1999).

95. Liaw, A. & Wiener, M. Classification and Regression by randomForest. **2**, (2002).

96. Amaratunga, D., Cabrera, J. & Lee, Y.-S. Enriched random forests. *Bioinformatics* **24**, 2010–2014 (2008).

97. Rybicki, E. The classification of organisms at the edge of life, or problems with virus systematics. *S. Afr. J. Sci.* **86**, 182–186 (1990).

98. S, M.-C., MK, J. & AM, M. COVID-19 and globalization. *One Heal. (Amsterdam,*

Netherlands) **9**, (2020).

99.   McKee, M. & Stuckler, D. If the world fails to protect the economy, COVID-19 will damage health not just now but also in the future. *Nat. Med. 2020 265* **26**, 640–642 (2020).

100.   Wei, S. *et al.* Field-deployable, rapid diagnostic testing of saliva for SARS-CoV-2. *Sci. Reports 2021 111* **11**, 1–9 (2021).

101.   Xun, G., Lane, S. T., Petrov, V. A., Pepa, B. E. & Zhao, H. A rapid, accurate, scalable, and portable testing system for COVID-19 diagnosis. *Nat. Commun. 2021 121* **12**, 1–9 (2021).

102.   Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565–574 (2020).

103.   Park, M., Won, J., Choi, B. Y. & Lee, C. J. Optimization of primer sets and detection protocols for SARS-CoV-2 of coronavirus disease 2019 (COVID-19) using PCR and real-time PCR. *Exp. Mol. Med.* **52**, 963–977 (2020).

104.   Bohn, M. K. *et al.* Molecular, serological, and biochemical diagnosis and monitoring of COVID-19: IFCC taskforce evaluation of the latest evidence. *Clin. Chem. Lab. Med.* **58**, 1037–1052 (2020).

105.   Ji, T. *et al.* Detection of COVID-19: A review of the current literature and future perspectives. *Biosens. Bioelectron.* **166**, (2020).

106.   Antibody, Antigen and Molecular Tests for COVID-19. *JOHNS HOPKINS CENTER FOR HEALTH SECURITY* (2021).

107.   WE, H. *et al.* RT-LAMP for rapid diagnosis of coronavirus SARS-CoV-2. *Microb. Biotechnol.* **13**, 950–961 (2020).

108.   Li, E., Larson, A., Kothari, A. & Prakash, M. Handyfuge-LAMP: low-cost and electricity-free centrifugation for isothermal SARS-CoV-2 detection in saliva. *medRxiv* 2020.06.30.20143255 (2020) doi:10.1101/2020.06.30.20143255.

109.   Pilarowski, G. *et al.* Field performance and public health response using the BinaxNOWTM Rapid SARS-CoV-2 antigen detection assay during community-based testing. *Clin Infect Dis* 1–27 (2020).

110.   Krüger, L. J. *et al.* Evaluation of the accuracy, ease of use and limit of detection of novel, rapid, antigen-detecting point-of-care diagnostics for SARS-CoV-2. *medRxiv* (2020) doi:10.1101/2020.10.01.20203836.

111.   Kevadiya, B. D. *et al.* Diagnostics for SARS-CoV-2 infections. *Nat. Mater.* **20**, 593–605 (2021).

112.   Ji, T., Liu, Z., Wang, G., Guo, X. & Akbar, S. Detection of COVID-19: A review of the current literature and future perspectives. *Biosens Bioelectron* **166**, 112455 (2020).

113.   Vandenberg, O., Martiny, D., Rochas, O., van Belkum, A. & Kozlakidis, Z. Considerations for diagnostic COVID-19 tests. *Nat. Rev. Microbiol.* (2020) doi:10.1038/s41579-020-00461-z.

114.   Sin, M. N. et al. Microbiological Guidelines for Food. *Cent. Food Saf.* **2014**, 1–38 (2014).

115.   Alpdagtas, S. *et al.* Evaluation of current diagnostic methods for COVID-19. *APL Bioeng.* **4**, 41506 (2020).

116.   Chen, Q. *et al.* Diagnostic technologies for COVID-19: a review. *RSC Adv.* **10**, 35257–35264 (2020).

117.   B, G. *et al.* Review of analytical performance of COVID-19 detection methods. *Anal. Bioanal. Chem.* **413**, 35–48 (2021).

118. Asghari, A. *et al.* Fast, accurate, point-of-care COVID-19 pandemic diagnosis enabled through advanced lab-on-chip optical biosensors: Opportunities and challenges. *Appl. Phys. Rev.* **8**, 031313 (2021).

119. Ahmad Bhutta, Z. *et al.* Nanophotonic biosensors for point-of-care COVID-19 diagnostics and coronavirus surveillance. *J. Phys. Photonics* **3**, 011002 (2021).

120. Goswami, N. *et al.* Label-free SARS-CoV-2 detection and classification using phase imaging with computational specificity. *Light Sci. Appl. 2021 101* **10**, 1–12 (2021).

121. Wyllie, A. L. *et al.* Saliva is more sensitive for SARS-CoV-2 detection in COVID-19 patients than nasopharyngeal swabs. *medRxiv* (2020) doi:10.1101/2020.04.16.20067835.

122. Li, Y. *et al.* Saliva is a non-negligible factor in the spread of COVID-19. *Mol. Oral Microbiol.* **35**, 141–145 (2020).

123. Sapkota, D., Thapa, S. B., Hasséus, B. & Jensen, J. L. Saliva testing for COVID-19? *Br. Dent. J.* **228**, 658–659 (2020).

124. Zamora, J. L. R. & Aguilar, H. C. Flow virometry as a tool to study viruses. *Methods* vols 134–135 87–97 (2018).

125. Reverberi, R. & Reverberi, L. Factors affecting the antigen-antibody reaction. *Blood Transfus.* **5**, 227–240 (2007).

126. Batéjat, C., Grassin, Q., Manuguerra, J.-C. & Leclercq, I. Heat inactivation of the severe acute respiratory syndrome coronavirus 2. *J. Biosaf. Biosecurity* **3**, 1–3 (2021).

127. Ward, G., Simpson, A., Boscato, L. & Hickman, P. E. The investigation of interferences in immunoassay. *Clin. Biochem.* **50**, 1306–1311 (2017).

128. Armbruster, D. A. & Pry, T. Limit of blank, limit of detection and limit of quantitation. *Clin. Biochem. Rev.* **29 Suppl 1**, S49-52 (2008).

129. Shrivastava, A. & Gupta, V. Methods for the determination of limit of detection and limit of quantitation of the analytical methods. *Chronicles Young Sci.* **2**, 21 (2011).

130. Corman, V. M. *et al.* Comparison of seven commercial SARS-CoV-2 rapid point-of-care antigen tests : a single-centre laboratory evaluation study. *The Lancet Microbe* **2**, 311–319 (2021).

131. Medical Device Coordination Group. *MDCG 2021-21 Guidance on performance evaluation of SARS-CoV-2 in vitro diagnostic medical devices*. (2021).