

# Field-Flow Fractionation (FFF) Project Lab Book

---

Note:

The complete repository can be found on Github. I made this repo public for the purpose of sharing it with you, it does not contain any sensitive or confidential materials. The repo can be found at <https://github.com/CheffySins/NMI-Placement-2024>.

This lab book is dated 13 October, and the labbook accessed on github may have been updated since the time of you opening this (can be found in /Lab Notes/ Lab Book.md). The main reason I'm sharing this, is so that you have quick direct access to the references if you need them, as well as to the photos and handwritten notes I took throughout my placement, as not all of them have since made their way to this digital notebook.

From speaking to my supervisor (Mar-Dean), it will be beneficial to her (and NMI), if I also submit this lab book in its final form at the end of Week 13 alongside the project report, as she'll then use it as a stepping stone for improving traceability with the AF4 measurements and data analysis, and therefore, this labbook has been laid out with that in mind (hence some of the missing photos, which are still there for your reference and my grades). I'm happy to come in and discuss any of the above with you if required, as I'd be able to better explain what I'm trying to get across

Read the README.md file if you'd like a breakdown of what every folder is used for.

## Table of Contents

---

- [Field-Flow Fractionation \(FFF\) Project Lab Book](#)
  - [Table of Contents](#)
  - [Project Overview](#)
  - [Executive Summary](#)
  - [Introduction](#)
  - [Project Plan](#)
    - [Aims and Objectives](#)
    - [Timeline](#)
  - [High-Level Overview and Background](#)
    - [Field-Flow Fractionation \(FFF\) Overview](#)
    - [Asymmetric Flow Field-Flow Fractionation \(AF4\)](#)
    - [Key Concepts](#)
  - [FFF / AF4 Machine Technical Term and Component Definitions](#)
    - [Software](#)

- Hardware
- Eclipse Instrument Components
- DAWN Instrument Components
- Theory and Key Concepts
  - Hydrodynamic Radius Theory & Calculation[^c-4]
  - Molar Mass Theory and Calculation
- Data Analysis and Processing
  - Data Types
  - Python Code Overview
    - Useful Libraries
    - Importing Data
    - Despiking Algorithm
    - Peak Selection Algorithm
    - Data Processing Techniques
- Experiment and Lab Work
  - Data Comparisons and Cross-Validation
  - Coffee Riffing Experiment
  - PSL in NaCl preparation
  - Remainder of Wet Lab Work
- References and Annotated Bibliography
  - Footnotes

## Project Overview

---

references can be done using [^1] and [^1]: at the end of the doc

Project Title: Developing Software for Data Processing from Field-Flow Fractionation (FFF) / Asymmetric Flow Field-Flow Fractionation (AF4)

Project Description: This project involves developing software to process data from a Field-Flow Fractionation (FFF) instrument, which is used to derive the size distributions of particles in liquid suspensions. This coding-intensive project requires experience with Python or a similar programming language to effectively analyse, clean, and visualise experimental data.

The team at NMI currently use Astra8 Software to do data analysis using their AF4 machine. This software acts as a kind of 'black box', in that obtaining fundamental proof of the algorithms / code used to do data treatment is either extremely difficult to obtain, or not possible (realistically). Additionally, the reports generated by Astra often contain too much non-useful information.

Traceability is a foundational cornerstone of nanometrology (and metrology in general), and so, I've been tasked with creating an open source template that will do the required data treatment, while also giving greater flexibility to displaying only the data that is required for a specific kind of experiment or sequence.

The order of treatments done by the team currently using Astra is:

1. Despiking
2. Baseline Selection
3. Peak Definitions
4. Molar Mass & Radius from Light Scattering (LS)<sup>[1]</sup>
5. Hydrodynamic Radius  $R_h$  from LS
6. Distribution Analysis

My goal is to write code that will do each of the above, both in that order, and individually. Once I've got together all of the equations and procedures required, I will transfer the code to Stata (provided by MQ), and then port into OriginLabs, which is the software that NMI staff have access to.

Location & Supervisor: National Measurement Institute (NMI), Lindfield Laboratory. Supervised by Mar-Dean Du Plessis

## Executive Summary

---

This lab book documents the development of a Python-based data processing software for the Field-Flow Fractionation (FFF) instrument utilised by the Nanometrology team at the National Measurement Institute (NMI). The FFF instrument is critical for deriving particle size distributions in liquid suspensions, an important aspect in nanoparticle characterisation. The focus of this placement was on Asymmetric Flow Field-Flow Fractionation (AF4), a specialised form of FFF. The software automates data processing tasks such as despiking, baseline correction, peak detection, and calculation of hydrodynamic radius. Calculating molar mass may be too difficult without an RI detector, and I have not done that yet but will attempt to if I have the capacity at the end. Preliminary results indicate significant improvements in efficiency and repeatability, enhancing NMI's capabilities in nanometrology by improving measurement accuracy and reducing manual workload.

## Introduction

---

Field-Flow Fractionation (FFF) is a widely used technique in nanometrology for separating and analysing particles based on size, shape, and density. The technique is highly effective for characterising nanoparticles in liquid suspensions, providing detailed insights into critical properties such as particle size distribution. Accurate measurements are essential for research in fields such as material science, biotechnology, and pharmaceuticals.

At NMI, the Nanometrology team employs FFF instruments to support research and provide services for

nano particle characterisation. However, the raw data generated by these instruments requires extensive manual processing, which is both labour-intensive and prone to inconsistencies. This project aims to automate the data processing pipeline using Python to enhance accuracy, repeatability, and efficiency, ultimately adding value to NMI's capabilities by reducing human error and standardising data analysis.

The specific focus of this project was on Asymmetric Flow Field-Flow Fractionation (AF4), a variant of FFF that uses a unique channel design to separate particles based on size. AF4 is particularly suited for analysing a wide range of particle sizes, from proteins to larger nanoparticles, making it a powerful tool in the field of nanometrology.

## Project Plan

---

### Aims and Objectives

The primary goal of this project is to develop a robust software solution for processing data generated by the AF4 instrument, enabling accurate determination of particle size distributions in liquid suspensions. The specific objectives include:

- Developing Python-based software for data handling, including despiking, peak selection, baseline correction, and size distribution analysis.
- Automating key data processing tasks to reduce manual workload and standardise procedures.
- Enhancing the efficiency and repeatability of AF4 data analysis.

### Timeline

This is a brief timeline and the tasks referenced will be discussed further below.

Week	Task(s)
1	Choose a project and familiarise myself with the theory <ul style="list-style-type: none"><li>- Lab Safety Induction</li></ul>
2	<ul style="list-style-type: none"><li>- Prepare solution of isopropanol for AF4 machine to stay in over the weekend</li><li>- Read up on relevant sections of MD's thesis</li></ul>
3	<ul style="list-style-type: none"><li>- Attend NMI award ceremony (online)</li></ul>
4	<ul style="list-style-type: none"><li>- Work on fine tuning code for the project</li><li>- Attend High Voltage lab tour</li></ul>
5	<ul style="list-style-type: none"><li>- Analyse and compare experiment results between the DLS and AF4 machines to validate data</li><li>- Investigate inconsistencies with UV plot data from AF4</li><li>- Shadow MD in wet lab coffee racking experiment (client work)</li></ul>

6	- Problem solving and troubleshooting issues and deviations in MADLS and DLS experiments
7	Helped MD trouble shoot issued with pressure build-up in AF4
8 onward	Primarily worked on the code, and assisted where relevant with wet lab work, troubleshooting pressure builds upds, etc.

## High-Level Overview and Background

### Field-Flow Fractionation (FFF) Overview

Field-Flow Fractionation (FFF) is a technique that allows the separation, or fractionation, of particles and macromolecules based on differences in their size, shape, and density. Fractionation in this context refers to the process of separating components of a mixture so that particles of similar properties are grouped together. In FFF, a field is applied perpendicular to the direction of flow, creating a velocity gradient that enables the differentiation of particles.

Field-Flow Fractionation (FFF) is a separation technique that allows the fractionation of particles and macromolecules based on differences in their hydrodynamic sizes. It works by applying a perpendicular field that forces particles towards the accumulation wall, creating a velocity gradient across the channel.

- Types of FFF: Sedimentation, Thermal, Flow, Centrifugal and Electrical FFF.
- FFF Instrument Components: Channel, flow pump, detector (e.g., Light Scattering, UV Absorbance).
- Applications: Protein aggregation analysis, nanoparticle characterisation, and colloid stability studies.

### Asymmetric Flow Field-Flow Fractionation (AF4)

Asymmetric Flow Field-Flow Fractionation (AF4) is a specific type of FFF that utilises an asymmetric channel, where only one wall allows cross-flow. This asymmetry creates a parabolic flow profile that enables highly efficient size-based separation of particles. AF4 is particularly useful for separating and analysing a wide range of particle sizes, from proteins and polymers to larger nanoparticles.

- AF4 Channel Design: AF4 uses a semi-permeable membrane on one side of the channel, which allows the cross-flow to drive smaller particles closer to the accumulation wall while larger particles remain in faster-moving regions.
- Applications: AF4 is ideal for nanoparticle characterisation, protein aggregation studies, and the analysis of complex colloidal systems.

### Key Concepts

#### *Absorbance A*

Quantifies how much light is absorbed by a sample. This is crucial in UV-Vis spectroscopy for determining

concentrations based on the Beer-Lambert law.

### Autocorrelation Functions $g_2$ and $g_1$

Autocorrelation functions describe the relationship between fluctuations in the scattered light over time, and they are fundamental to Dynamic Light Scattering (DLS). In DLS, two related autocorrelation functions are used to obtain information about particle diffusion, size, and interactions within a medium:

Intensity Autocorrelation Function  $g_2(\tau)$ : This second-order autocorrelation function measures the correlation of light intensity fluctuations over time. It is directly related to the movement of particles through Brownian motion. As particles diffuse, the scattered light intensity varies, and  $g_2(t)$  captures these time-dependent changes, providing data for particle size distribution.

The function is defined as:

$$g_2(\tau) = \langle I(0)I(\tau) \rangle$$

Where:

- $I(0)$  represents the scattered light intensities at times 0 and  $\tau$ , and
- $\langle \cdot \rangle$  denotes a time average.

Electric Field Autocorrelation Function  $g_1(\tau)$ : This first-order autocorrelation function represents the correlation of the scattered electric field over time, which is more fundamental than  $g_2(\tau)$ .

For a dilute<sup>[2]</sup>, monodisperse nanoparticle solution,  $g_1(\tau)$ 's time delay may be calculated by the translational self-diffusion coefficient  $D$  and the scattering vector length  $q$ :

$$g_1(\tau) = e^{-q^2 D \tau}$$

The relationship between  $g_1(\tau)$  and  $g_2(\tau)$  is often expressed:

$$g_2(\tau) = 1 + \beta |g_1(\tau)|^2$$

Where:

- $\beta$  is the coherence factor, and
- $g_1(\tau)$  provides insight into the diffusion coefficient  $D$  and, subsequently, particle size through the Stokes-Einstein relationship.

### Exponential Decay $\Gamma$

Represents the rate of decay of the autocorrelation function in DLS and is associated with the particle's diffusion coefficient:

$$\Gamma = Dq^2$$

Where  $q$  is the scattering vector. This term provides insight into how quickly particles are diffusing, which

is directly related to their size.

Exponential Decay Fit, modelling  $g_2(\tau) - 1$  as:

$$g_2(\tau) - 1 = Ae^{(-2\Gamma\tau)}$$

### Coherence Factor $\beta$

The coherence factor  $\beta$  appears in the relationship between  $g_2$  and  $g_1$  and is an instrumental parameter related to the detection optics and the alignment of the measurement setup in DLS. It typically has a value between 0 and 1, where in an ideal apparatus it is equal to 1 (Pecora, 2000, p.125) and represents the extend of spatial coherence of the scattered light reaching the detector. The factor  $\beta$  depends on:

- Instrument alignment: Perfect alignment increases the value of  $\beta$ .
- Laser coherence: A coherent laser sources also contributes to a higher  $\beta$
- Detection configuration: The size and position of the detector aperture affect the spacial coherence of the signal.

The coherence factor thus influences the maximum value that  $g_2(\tau)$  can reach and helps to determine the quality of data obtained from DLS measurements.

Time Delay  $\tau$  Time delay (correlation time) is the time interval used in DLS to measure how  $g_2(\tau)$  changes over time. It represents the lag time between two intensity measurements, allowing for analysis of the time behaviour of particle motion. In DLS,  $\tau$  is varied systematically to create a range of time delays, which are then used to build  $g_2(\tau)$ , and the rate at which  $g_2(\tau)$  decays is directly related to the diffusion coefficient of the particles (and therefore their size).

### Beer-Lambert Law

The Beer-Lambert Law relates the absorbance of light to the properties of a material through the following equation:

$$A = \epsilon cl$$

Where:

- $A$  is the absorbance,
- $\epsilon$  is the molar extinction coefficient,
- $c$  is the concentration of the absorbing species, and
- $l$  is the path length of the sample cell.

This law is fundamental in UV-Vis spectroscopy for quantifying the concentration of analytes in solution. It assumes that the absorbance is directly proportional to both the concentration of the substance and the

path length. This linear relationship is valid for dilute solutions, where deviations due to high concentrations or chemical interactions are minimal. Typically, this law holds up to concentrations around 0.01 M, where beyond this limit, deviations occur due to:

- High concentration effects, where the solute molecules may interact or aggregate, changing their absorbance characteristics.
- Nonlinear absorption, the absorption of light may saturate at higher concentrations, especially in cases where self-absorption or re-emission occurs.

The exact concentration threshold depends on the chemical nature of the substance and the specific experimental conditions, but generally, it is best to keep concentrations under 0.01 M to maintain linearity in the Beer-Lambert law for accurate measurements.

### *Burchard Stockmayer Shape Factor*

This shape factor relates the radius of gyration  $r_g$  to the hydrodynamic radius  $r_h$  and provides information about the overall shape and structure of macromolecules:

$$\text{Shape Factor} = \frac{r_g}{r_h}$$

This factor can indicate whether a molecule has a more compact, spherical structure (lower values) or an elongated, rod-like structure (higher values). This is useful for characterising the three-dimensional shape of polymers and other macromolecules in solution.

### *Colloidal Systems*

Mixtures where small particles, typically ranging from 1 nm to 1  $\mu\text{m}$  in size, are dispersed throughout a continuous medium (liquid, gas or solid). These dispersed particles do not settle out over time due to gravitational forces and remain suspended due to Brownian motion or electrostatic interactions.

### *Detector Constant $K_{\text{detector}}(\theta)$*

A constant specific to the detector and angle  $\theta$ , used to relate the measured scattering intensity to the Rayleigh ratio.

### *Differential Refractive Index Increment $dn/dc$*

Represents the change in refractive index as the concentration of a solute changes. It is a key parameter in Multi-Angle Light Scattering (MALS) for calculating molecular weights.

### *Diffusion Coefficient $D$*

The diffusion coefficient,  $D$ , quantifies the rate at which particles move through a medium due to Brownian motion. It describes the particle's tendency to spread from regions of higher concentration to regions of lower concentration and is central to characterising the motion of particles in Dynamic Light Scattering (DLS).

Broadly, the diffusion coefficient can be described by Fick's first law

$$J = -D \frac{dC}{dx}$$

Where:

- $J$  is the flux of particles,
- $D$  is the diffusion coefficient,
- $C$  is the concentration, and
- $x$  is the position.

In DLS, the diffusion coefficient (often referred to as the translational self-diffusion coefficient) specifically describes the movement of individual particles in a homogenous solution, independent of concentration gradients or external forces. The term 'self-diffusion' emphasises that the measurement pertains to the intrinsic, random motion of particles due to thermal energy, without any net movement driven by external gradients. In DLS,  $D$  is related to the decay rate  $\Gamma$  of the electric field autocorrelation function  $g_1(t)$  and the scattering vector  $q$  by:

$$\Gamma = Dq^2$$

Where  $q$  is the scattering vector

For spherical particles,  $D$  is related to the hydrodynamic radius  $r_h$  through the Stokes-Einstein equation, which is commonly used in DLS to derive particle size from diffusion measurements, assuming that the particles are spherical and undergoing Brownian motion.

### *Dynamic Light Scattering (DLS) and Multi-Angle Light Scattering (MALS)*

Dynamic Light Scattering (DLS) and Multi-Angle Light Scattering (MALS) are two light scattering techniques widely used in particle and molecular characterisation.

- DLS, also known as quasi-elastic light scattering (QELS), measures the time-dependent fluctuations in scattered light intensity caused by the Brownian motion of particles in suspension. By analysing these fluctuations, DLS calculates the diffusion coefficient and, subsequently, the hydrodynamic radius of particles. This technique is especially useful for determining particle sizes in the nanometer range, such as proteins polymers, and nanoparticles.
- MALS measures the angular dependence of scattered light to determine the molecular weight and radius of gyration of macromolecules in solution. Unlike DLS, which is limited to particle size, MALS can characterise a wider range of macromolecular properties, including shape, structure, and conformation, making it essential in fields like polymer science and biophysics.

### *Form Factor $P(\theta)$*

Represents how the scattering intensity depends on the particle shape and size at a particular angle  $\theta$ . It is

critical in interpreting the angular dependence of scattered light.

### Hydrodynamic Diameter $d_h$

This value represents twice the hydrodynamic radius and is given by  $d_h = 2r_h$ . It provides a direct measure of particle size in relation to how it behaves hydrodynamically in the surrounding medium, which is essential in determining the full scale of particle movement and interaction within the medium.

### Hydrodynamic Radius $r_h$

Represents the effective radius of a particle in suspension as it experiences hydrodynamic drag, a measure that reflects how a particle moves through a medium under an external force. This is particularly useful in characterising particles in a fluid and is calculated using the Stokes-Einstein equation

$$r_h = \frac{k_B T}{6\pi\eta D}$$

Where:

- $k_B$  is the Boltzmann constant,
- $T$  is the absolute temperature,
- $\eta$  is the viscosity of the medium, and
- $D$  is the diffusion coefficient

This relationship highlights that  $r_h$  depends on the temperature and viscosity of the medium as well as the diffusion behaviour of the particle.

### Molar Extinction Coefficient $\epsilon$

A measure of how strongly a chemical species absorbs light at a particular wavelength, allowing the calculation of concentration from absorbance measurements.

### Mie Theory

Describes the scattering of electromagnetic waves by spherical particles, applicable when the particle diameter is comparable to the wavelength of the incident light. Developed by Gustav Mie, the theory is an extension of Rayleigh scattering to larger particles. Mie theory provides the framework to calculate scattering intensity as a function of the particle size, refractive index, and wavelength of the incident light, and is essential for interpreting light scattering from particles such as aerosols, cells, and other systems. The theory has applications in fields like medical imaging, material science and atmospheric science.

### Optical Constant $K$

A constant used in light scattering calculations that incorporates the refractive index and wavelength of light. It is vital for interpreting scattering data accurately.

### Osmotic Pressure

This was not really used during any of the project, I just found it interesting and read more about it once I saw it mentioned next to the Virial Coefficient sections in the equations, which describe aggregation in

particles, a core concept in nanometrology.

Osmotic pressure is the pressure required to prevent the flow of a solvent into a solution through a semipermeable membrane. It arises when a solvent moves from a region of lower solute concentration (pure solvent) to a region of higher solute concentration (solution) to equalise concentrations on both sides of the membrane. This flow is driven by osmotic forces, which are rooted in the tendency of systems to reach equilibrium.

In a solution, osmotic pressure ( $\Pi$ ) can be described by van't Hoff's equation for dilute solutions

$$\Pi = iMRT$$

Where:

- $i$  is the van't Hoff factor (number of particles per formula unit),
- $M$  is the molarity of the solution,
- $R$  is the gas constant, and
- $T$  is the absolute temperature.

Osmotic pressure is a key concept in biological systems, such as in maintaining cell turgor, and in fields like chemistry and environmental science, where it affects processes like water purification and desalination.

### *Radius of Gyration $r_g$*

A measure of the distribution of a particle's mass around its center of mass and is especially relevant for larger molecules (typically over 10nm). Using multi-angle light scattering (MALS),  $r_g$  can be determined by examining the scattering intensity  $I(q)$  relative to the scattering vector  $q$ :

$$R_g^2 = \frac{I(q)}{q^2}$$

Where  $q$  is the scattering vector.

The radius of gyration provides insight into the particle's structural shape, indicating how mass is dispersed from the particle's center, which can be crucial for determining molecular configurations.

### *Rayleigh Ratio $R_{std}(\theta)$*

Describes the intensity of scattered light from a standard sample at a given angle  $\theta$  and is used to calibrate the scattering intensity for quantitative analysis.

### *Rayleigh Scattering*

A type of scattering that occurs when the particles are much smaller than the wavelength of light. It describes the relationship between scattered light and particle size, often utilised in Multi-Angle Light Scattering (MALS) or macromolecular characterisation.

## *Recovery % (from UV Data)*

Used to quantify the amount of sample successfully eluted and detected. It is calculated as follow:

$$\text{Recovery \%} = \left( \frac{\text{Amount detected}}{\text{Initial amount}} \right) \times 100$$

This is particularly useful in quality control and process optimisation, as it provides a simple measure of how much of the sample initially placed in the system is eventually recovered and measured.

## *Refractive Index $n$*

Describes how light propagates through a medium. In light scattering, it is essential for calculating the scattering vector  $q$  and understanding the interaction between light and particles within the medium.

## *Scattering Vector $q$*

The scattering vector describes the difference between the incoming and outgoing light wave vectors during scattering. It is defined as:

$$q = \frac{4\pi n}{\lambda} \sin\left(\frac{\theta}{2}\right)$$

Where:

- $n$  is the refractive index,
- $\lambda$  is the wavelength of the incident light, and
- $\theta$  is the scattering angle.

The scattering vector is crucial in determining structural information from scattering data as it relates to the size and shape of the particles.

## *Second Virial Coefficient $B_2$* [3]

The second virial coefficient,  $B_2$ , is a parameter that describes the interaction between pairs of molecules in solution. It appears in the virial expansion of the osmotic pressure and in equations for light scattering. Physically,  $B_2$  reflects how attractive or repulsive forces between molecules affect their behaviour in a solution. It is included in the following expansion for osmotic pressure  $\Pi$ :

$$\Pi = \frac{RTc}{M} (1 + B_2 c + B_3 c^2 + \dots)$$

- $R$  is the gas constant,
- $T$  is the absolute temperature,
- $c$  is the concentration,
- $M$  is the molar mass, and

- $B_3, B_4$ , etc., are higher-order virial coefficients.

In light scattering,  $B_2$  provides insight into the molecular interactions:

- If  $B_2$  is positive, it indicates repulsive interactions between particles, leading to a more expanded structure in solution.
- If  $B_2$  is negative, it suggests attractive interactions, which can lead to aggregation or clustering.

$B_2$  is therefore a key factor in understanding the stability and aggregation behaviours and tendencies of macromolecules in solution.

### Spectroscopy

Spectroscopy is the study of the interaction between electromagnetic radiation and matter. It is a cornerstone tool in analytical chemistry and physics for determining the composition, structure and properties of materials. By analysing how different wavelengths of light are absorbed, emitted or scattered by a substance, scientists can infer information about its molecular makeup and behaviour. Spectroscopic techniques are applied across various regions of the electromagnetic spectrum, including ultraviolet (UV), visible (Vis), infrared (IR), and X-ray, each showcasing novel insights depending on the sample and the properties being investigated.

### Stokes-Einstein Equation

The Stokes-Einstein equation describes the relationship between the diffusion coefficient  $D$  of a spherical particle and its hydrodynamic radius  $r_h$ :

$$r_h = \frac{k_B T}{6\pi\eta D}$$

- $k_B$  is the Boltzmann constant,
- $T$  is the absolute temperature,
- $\eta$  is the viscosity of the medium, and
- $D$  is the diffusion coefficient.

This equation is foundational in Dynamic Light Scattering (DLS) as it allows for the determination of particle size based on diffusion measurements. The equation assumes that the particles are spherical and are diffusing through a medium under Brownian motion. The Stokes-Einstein equation gives crucial insights into the relationship between a particle's size, temperature, and the fluid's viscosity, enabling researchers to characterise nanoparticles, proteins, and other colloidal systems.

### UV-Vis Spectroscopy

UV-Vis spectroscopy is a type of absorption spectroscopy that focuses on the ultraviolet and visible regions of the electromagnetic spectrum. It measures how much UV or visible light a sample absorbs, often providing information about the electronic transitions in molecules. This technique is commonly used for determining the

concentration of solutions through the Beer-Lambert law, which relates the absorbance of light to the properties of the material. UV-Vis spectroscopy is invaluable for analysing compounds with conjugated electron systems, such as organic molecules and biomolecules, and is widely used in fields like biochemistry, environmental science, and pharmacology.

### Viscosity $\eta$

The measure of a fluid's resistance to flow . It is a crucial parameter in the Stokes-Einstein equation for calculating the hydrodynamic radius and impacts how particles move through the fluid.

### Zimm Theory

Zimm theory, often used in conjunction with static light scattering, describes the molecular weight, size, and shape of polymers in solution. The theory, developed by Bruno Zimm, includes a model that accounts for inter-molecular interactions and uses the concept of the Zimm plot, which plots the inverse of scattered light intensity against the scattering angle. The slope and intercept of the Zimm plot provide important parameters such as the molecular weight, radius of gyration, and second virial coefficient. Zimm modelling helps to understand polymer structure and dynamics, making it a fundamental tool in polymer chemistry and biochemistry.

## FFF / AF4 Machine Technical Term and Component Definitions

### Software

Term	Description
Elution	In FFF, elution refers to the process where separated particles exit the channel. Smaller particles typically elute later due to slower diffusion rates in AF4 systems.
Focus	A step where the sample is concentrated at a specific point in the channel by applying flows from opposite directions, positioning the particles before separation.
Focus Inject	A combined step where the sample is injected and simultaneously focused. The sample is introduced into the channel and pushed to the focus area, preparing it for separation.
Elution Inject	Involves injecting the sample into the channel specifically for the elution phase. It follows the focus phase and marks the start of particle separation and elution.
Cross Flow	In AF4, cross flow is a perpendicular flow applied across the channel width. It pushes particles toward the channel walls, creating a size-based separation.
Channel Flow	The main flow that moves the sample down the length of the channel toward the outlet, carrying particles through the channel for separation by cross flow.
Outlet Flow	The final flow that directs particles out of the channel after separation. Particle properties, such as size and shape, affect the time taken to reach the outlet.

## Hardware

### Eclipse Instrument Components

Component	Description	Source
Membranes	Made from polyethersulfone or regenerated cellulose with molecular weight cut-offs ranging between 2-30 kDa. Used in channels for filtration and separation. Custom membranes can be cut for specific needs.	TN6006D Eclipse and Vision Handbook, Eclipse Channel Overview and Membrane Installation
Channel Modules	Channels available in various sizes (Short, Long, Dispersion Inlet, Semi-Prep) and compatible with both aqueous and organic solvents. Each channel can regulate temperature up to 50°C. Different spacers ensure proper channel sealing.	TN6006D Eclipse and Vision Handbook, Channel Design Section
Inline Filters	Installed on the pump, cross flow, and dilution control module (DCM) ports. These filters help prevent particulate contamination and must be replaced periodically for optimal performance.	TN6006D Eclipse and Vision Handbook, Bulkhead Frit Filter and Inline Filter Assemblies
Cross Flow and DCM Ports	Specialized flow ports for cross-flow separation and dilution control to enhance signal concentration without sacrificing resolution.	TN6006D Eclipse and Vision Handbook, Side-Panel Connections
Temperature Regulator	Maintains a controlled environment for the channel and fluid pathways, essential for consistent analysis results.	TN6006D Eclipse and Vision Handbook, Channel Temperature Regulator Connection
Fluid Leak Sensors	Positioned at multiple points for safety, these sensors detect leaks and automatically activate alarms, especially important for preventing damage when using solvents.	TN6006D Eclipse and Vision Handbook, Eclipse Liquid and Vapor Leak Detection

### DAWN Instrument Components

Component	Description	Source
	Designed for minimal volume to allow small sample measurements. Composed of fused silica, optimal for refractive index measurements.	DAWN User's

Flow Cell	The cell assembly includes glass windows aligned with the laser to measure scattered light.	Guide, Chapter 2, Flow Cell Design
Interference Filters	Filters attached to specific photodiodes to reduce background noise and improve measurement accuracy, particularly for samples that fluoresce under laser light.	DAWN User's Guide, Appendix E, Installing Interference Filters
WyattQELS Module	An embedded Dynamic Light Scattering module for real-time measurements of hydrodynamic radius, useful in both batch and online (in-line) configurations.	DAWN User's Guide, Chapter 1, Instrument Options and Accessories
Laser and Detectors	The DAWN's laser measures at 18 angles simultaneously, which enhances the detection of light scattering for macromolecular analysis. The detectors are arranged strategically around the laser to capture a broad range of angles.	DAWN User's Guide, Chapter 2, Detector Placement
Inline Filters	Similar to the Eclipse, DAWN uses inline filters for both aqueous and organic applications. These filters prevent particulates from interfering with light scattering measurements.	DAWN User's Guide, Chapter 1, Inline Filter Kits
COMET Module <sup>[4]</sup>	Ultrasonic transducer installed in the DAWN instrument that automatically cleans the flow cell to prevent particle buildup.	DAWN User's Guide, Instrument Options and Accessories

## Theory and Key Concepts

### Hydrodynamic Radius Theory & Calculation<sup>[5]</sup>

The calculation of  $r_h$  for this project is done through the following process:

1. Calculate Intensity Autocorrelation Function  $g_2(\tau)$ :

$$g_2(\tau) = \frac{\langle I(t) \cdot I(t+\tau) \rangle}{\langle I(t) \rangle^2}$$

Where:

- asd asdjasdhsajdh:
- $I(t)$  is the intensity at time  $t$
- $\langle I \rangle$  is the mean intensity, and

- $\tau$  is the delay time

2. Note Electric Field Autocorrelation Function  $g_1(\tau)$ :

$$g_1(\tau) = e^{(-\Gamma\tau)}$$

Where  $\Gamma$  is the decay rate.

3. Relate  $g_2(\tau)$  to  $g_1(\tau)$ :

$$g_2(\tau) = 1 + \beta|g_1(\tau)|^2$$

Where  $\beta$  is the experimental coherence factor, which is instrument / set-up specific (approximated as 1).

1. Exponential Decay Fit, modelling  $g_2(\tau) - 1$  as:

$$g_2(\tau) - 1 = Ae^{(-2\Gamma\tau)}$$

Where  $A$  is the amplitude

5. Calculate the Scattering Vector  $q$  for each detector angle  $\theta$ :

$$q = \frac{4\pi n}{\lambda} \sin\left(\frac{\theta}{2}\right)$$

Where:

- $n$  is the refractive index of the solvent,
- $\lambda$  is the laser wavelength (nm), and
- $\theta$  is the scattering angle in degrees.

6. Calculate the Diffusion Coefficient  $D$ , using its relationship to  $\Gamma$ :

$$\Gamma = Dq^2 \Rightarrow D = \frac{\Gamma}{q^2}$$

7. Use the Stokes-Einstein Equation to Calculate the  $r_h$ :

$$D = \frac{k_B T}{6\pi\eta r_h} \Rightarrow r_h = \frac{k_B T}{6\pi\eta D}$$

Where:

- $k_B$  is the Boltzmann constant,
- $T$  is the absolute temperature, and
- $\eta$  is the viscosity of the solvent.

Note: different detector angles provide multiple values for  $q$ , improving the accuracy of the calculated

hydrodynamic radius by reducing uncertainty through averaging

## Molar Mass Theory and Calculation

This section requires many assumed values, as well as a refractometer for  $dn/dc$  values, which is not currently available at NMI<sup>[6]</sup>.

Where a sample with the required known values is used<sup>[5:1]</sup>, Molar Mass ( $M$ ) can be obtained using the following steps:

1. Calibrate Detectors:

- Using a standard sample with known Rayleigh Ratio  $R_{std}(\theta)$  and concentration  $c_{std}$ .
- Measure detector voltages  $V_{std}(\theta)$  for the standard

2. Calculate the Detector Constant  $K_{detector}$ :

$$K_{detector} = \frac{R_{std}(\theta)}{V_{std}(\theta)}$$

1. Compute Sample Rayleigh Ratios:

$$R_{std}(\theta) = K_{detector} \cdot V_{sample}(\theta)$$

1. Determine Sample Concentration  $c$ :

$$A = \epsilon cl \Rightarrow c = \frac{A}{\epsilon l}$$

Where:

- $A$  is the absorbance from the UV data.
- $\epsilon$  is the molar extinction coefficient ( $L \cdot mol^{-1} \cdot cm^{-1}$ ), and
- $l$  is the machine cell path length, in cm.

1. Calculate the Optical Constant  $K$ :

$$K = \frac{4\pi^2 n^2 (dn/dc)^2}{N_A \lambda^4}$$

Where:

- $n$  is the refractive index of the solvent,
- $dn/dc$  is the refractive index increment ( $mL \cdot g^{-1}$ ),
- $N_A$  is Avogadro's number, and
- $\lambda$  is the wavelength of the laser in cm.

## 1. Use the Rayleigh Equation to Calculate $M$ <sup>[7]</sup>:

- For Small Particles (Rayleigh Scattering):

$$R(\theta) = KcM \Rightarrow M = \frac{R(\theta)}{Kc}$$

- For Larger Particles, incorporate the form factor  $P(\theta)$ :

$$R(\theta) = KcMP(\theta)$$

Where  $P(\theta) = e^{-\frac{(qR_g)^2}{3}}$  is the form factor for spherical particles, and  $r_g$  is the radius of gyration

Note: This calculation assumes isotropic particles, and deviations may occur for non-spherical particles, which could lead to inaccuracies. Care should be taken to validate assumptions about particle shape during analysis.

## Data Analysis and Processing

### Data Types

- Time Data: Represents elapsed time during the fractionation process.
- Raw Light Scattering Data: Unprocessed measurement of scattered light intensity detected during a DLS / MALS experiment. Represents intensity of light scattered by particles as a function of time.
- UV Absorbance Data: Measurement of the amount of UV light absorbed by a sample over a range of wavelengths. Useful for concentration analysing structural properties of the particles.
- Count rate: A measure of the frequency of detected photon events. Represents the number of photons scattered by the particles and detected by the photodetector per second. The count rate is directly related to the intensity of scattered light.
  - A higher count rate generally indicates larger particles or higher concentrations as more photons are scattered (meaning increased LS intensity).
  - Stable count rate suggests consistent scattering and uniform particle distribution, whereas a fluctuating one may mean aggregation or the presence of impurities.

### Python Code Overview

Note this code is not final, and will be amended in both the report and the final notebook that is sent to NMI at the end of Week 13.

### Useful Libraries

The list of libraries used is not yet complete, as the code is not yet final. This section is a work in progress and has taken a lower priority on my list of things to have completed.

I have used the following libraries:

- pandas for data manipulation and analysis, particularly with data structures like dataframes, which are great for handling tabular data.
- numpy for numerical computations in Python, providing support for large, multi-dimensional arrays and matrices.
- matplotlib.pyplot, which is a plotting library used for creating static, interactive, and animated visualizations in Python, particularly for 2D plots and graphs.
- scipy.signal.medfilt, a function within the SciPy library's signal processing module, used for applying a median filter to a signal, which is helpful in reducing noise in data.

## Importing Data

The below code all imports data using the pandas module, using heading columns from the ASTRA exports

```
import pandas as pd
data = pd.read_csv('747.csv') # In origin, this can be made to be whatever csv is inserted.
# These are purely examples.
time = data['time (min)']
signal_intensity = data['raw light scattering data: detector voltage (V) detector 11']
absorbance_UV1 = data['UV absorbance data: absorbance (AU) channel 1']
absorbance_UV2 = data['UV absorbance data: absorbance (AU) channel 2']
```

## Despiking Algorithm

Despiking is the process of identifying and removing extreme deviations (spikes) in data. Spikes in data are typically due to noise or external anomalies and do not represent the underlying signal. Example can be seen in the raw data plots in the Practical Notebook. There are three different despiking algorithms that I investigated and tested;

1. Smoothing / Rolling Mean Method
2. Z-Score Method
3. Median Filter Method

## Z-Score Method Equation

A z-score / Standard Score describes a value's relationship to the mean of a group of values. It indicates how many standard deviations a data point is from the mean of the dataset. The formula for calculating the z-score of a value  $x$  is:

$$z = \frac{x - \mu}{\sigma}$$

- $x$  is the value being evaluated.

- $\mu$  is the mean of the dataset.
- $\sigma$  is the standard deviation of the dataset.

Interpretation:

- A z-score of 0 means the data point is exactly at the mean.
- A positive z-score indicates the data point is above the mean.
- A negative z-score indicates the data point is below the mean.
- For example, a z-score of 2 means the data point is 2 standard deviations above the mean, while a z-score of -1.5 means it is 1.5 standard deviations below the mean.

Description Table

Method	Description	Applications	Notes
Rolling Mean	A technique where each data point in a time series is replaced by the average of the neighboring points within a defined window. This method smooths the data by reducing the impact of random noise and spikes.	Useful when the goal is to smooth out short-term fluctuations and highlight longer-term trends or cycles in the data. It's commonly used in any context where gradual trends are more important than short-lived variations.	While Rolling mean is useful for identifying overall trends, it can smooth out significant but brief variations (potentially losing important data features). Additionally, rolling mean is not very reliable at boundaries due to fewer points being available.
Z-Score	Involves calculating the standard score (Z-score) of each data point, which measures how many standard deviations a point is from the mean. Data points with Z-scores exceeding a certain threshold are considered outliers or spikes and can be replaced or removed.	best suited for situations where outliers are expected to be rare and significantly different from the rest of the data, such as in sensor readings or quality control processes. It's also useful when the data is approximately normally distributed.	Useful for identifying significant outliers, it assumes that the data is normally distributed, which is not always the case.
Median Filter	A median filter replaces each data point with the median of its neighbors within a specified window. Unlike the rolling mean, the median filter is robust to outliers because the	Effective in applications where spikes are frequent but not indicative of true data trends, such as in image processing, environmental sensor data, or any scenario where random,	Particularly useful to preserve sharp features / edges within the data.

median is less affected by extreme values.

sharp noise needs to be removed without distorting the underlying signal.

below markdown python formatting for reference

```
from scipy.signal import medfilt # Libraries used
import matplotlib.pyplot as plt

kernel_size = 5 # Window size
median_signal_intensity = medfilt(signal_intensity,kernel_size) # Median Filter

# Plot the original and despiked data
fig3 = plt.figure(figsize=(15, 6))
plt.plot(time,signal_intensity, label='Original Data')
plt.plot(time,median_signal_intensity,'r--' , label='Despiked Data')
plt.ylabel('Detector 11 Voltage (V)')
plt.xlabel('Time (min)')
plt.title('Despiking Using Median Filter')
plt.legend()
plt.show()
```

## Peak Selection Algorithm

```
import numpy as np
import matplotlib.pyplot as plt

# Automatic peak detection using thresholds
threshold = np.max(voltage_corrected) * 0.05 # Adjust threshold as needed, currently at 5%
peak_indices = np.where(voltage_corrected > threshold)[0]

# Separate peaks based on gaps
gap_threshold = 10 # Set a threshold for the gap
peak_groups = []
current_group = [peak_indices[0]]

for i in range(1, len(peak_indices)):
    if peak_indices[i] - peak_indices[i - 1] > gap_threshold:
        peak_groups.append(current_group)
        current_group = [peak_indices[i]]
    else:
        current_group.append(peak_indices[i])
peak_groups.append(current_group) # Append the last group

# Plot the voltage signal with each peak in a different color
```

```

plt.figure(figsize=(12, 6))
plt.plot(time, voltage_corrected, label='Corrected Voltage', color='gray')

# Plot each peak group with a different color
colors = plt.cm.get_cmap('tab10', len(peak_groups))

for i, group in enumerate(peak_groups):
    time_peak = time[group]
    voltage_peak = voltage_corrected[group]
    plt.plot(time_peak, voltage_peak, '.', color=colors(i), label=f'Peak {i+1}')

plt.xlabel('Time (min)')
plt.ylabel('Corrected Voltage (V)')
plt.title('Peak Selection with Distinct Peaks')
plt.legend()
plt.show()

# Store peaks in a dictionary for further use
peak_data = {f'Peak {i+1}': {'time': time[groups], 'voltage': voltage_corrected[groups]}}
    for i, groups in enumerate(peak_groups)}

# Peak data points can be accessed using the peak_data directory

```

## Data Processing Techniques

- Baseline Correction: Removes baseline noise to provide a more accurate representation of peak data.
- Peak Integration: Calculates the area under peaks to estimate size and concentration.

## Experiment and Lab Work

---

All references to Mar-Dean are abbreviated as MD.

All images references are either in lab books 1/2 or in the lab work folder. Footnotes added.

## Data Comparisons and Cross-Validation

30 August 2024

Assisted MD with comparing Polystyrene Latex (PSL) in SDS data across the two machines. The data for the 100nm particles is much nicer than the 20nm data when looking at the UV absorption data, illustrated in Lab Notes 2 PDF<sup>[8]</sup>.

The primary issue identified is that the UV plots follow a different pattern / have a different shape to the DLS data.

Potential solutions to this issue were:

- Trying to emulate tested BSA method for use with PSL. Since BSA is ~5nm, it 'should' work on larger particles.
- Looking through the Wyatt handbook for 50nm particles and copying their methods to check if that works

The Wyatt handbook for PSL with long channel (what NMI has) had the following method:

Step	Duration (min)	Cross-flow Start (mL/min)	Cross-flow end (mL/min)
1. Elution	1	0	1
2. Elution	2	1	1
3. Focus	1	1	1
4. Focus Inject	3	1	1
5. Focus	6	1	1
6. Elution	30	0.5	0.5
7. Elution	5	0.5	0.00
8. Elution Inject	10	0.00	0.00

Step 1 was added by MD, as the machine tends to produce better results when allowed to 'ramp up'.

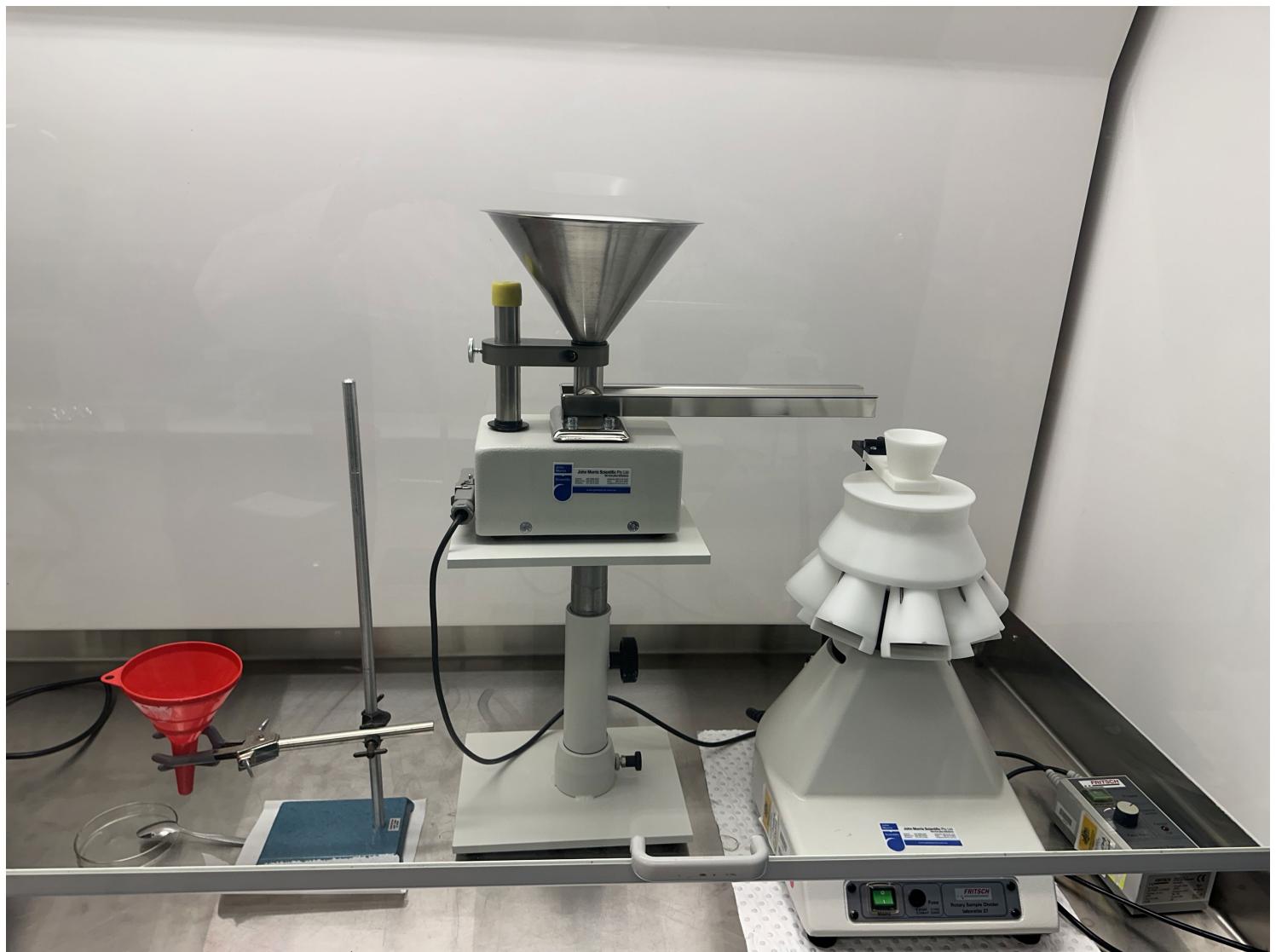
## Coffee Riffing Experiment

### Aim

We want to get a true representation of a bulk sample, removing bias that is added during things steps such as transport (e.g. truck driving the sample, which vibrates the sample and causes coffee to be dispersed by size).

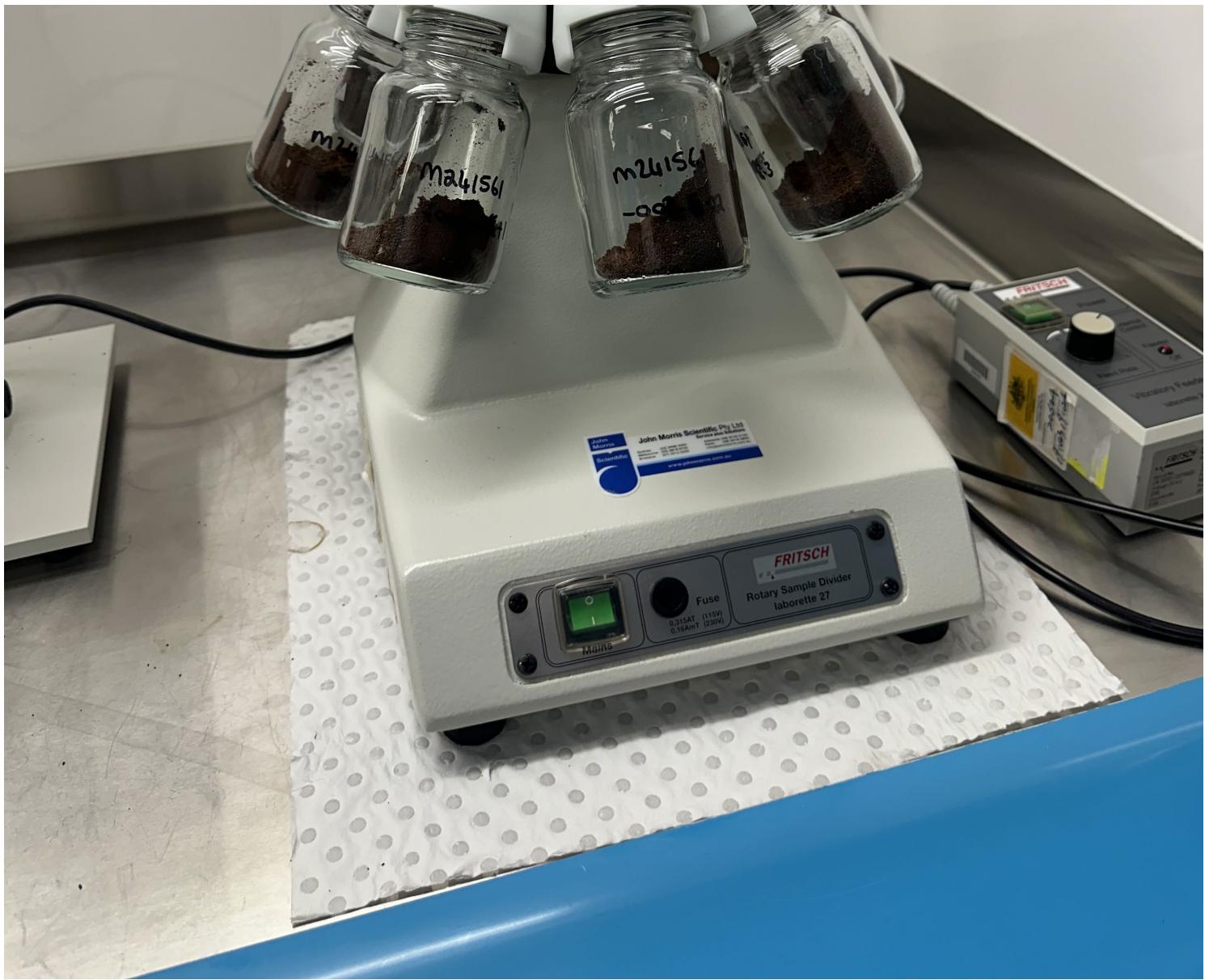
### Steps

1. Label everything
2. Weigh the coffee
3. Weigh the jars
4. Line up the jars with riffler and attach all 8
  - Riffler before jar attachments



- Riffler after attaching the jars





5. Add sample by the spoonful to the funnel<sup>[9]</sup>

Coffee splits over the channel due to vibrations and travels along it into the spinning funnel, separating into 1 of 8 jars. The idea is, by spinning, the jars collect more even distributions of the coffee particles that are bouncing along the channel.

Sometimes, coffee builds up after going into the separator (before entering jars), potentially due to:

- Coffee heats up (from channel), or
- Adding the coffee too fast, causing it to clump.

6. Brush any remaining sample into channel

7. Turn off the machine after all possible sample is riffled

8. Remove jars from rotator

9. Weigh full jars
10. Pour into plastic jars
11. Choose 1 sample to split into a further 8 vials
12. Repeat experiment
13. Add to small vials
14. Weight vials

## PSL in NaCl preparation

06 September 2024

### *Image References*

Lab Notes 2 & 1 PDFs

Started the day in lab by placing AF4 machine in isopropanol for the weekend.

The aim for today was to:

- Gain a better understanding of the DLS machine & DLS techniques.
- Assist in preparing 10mMol NaCl solution

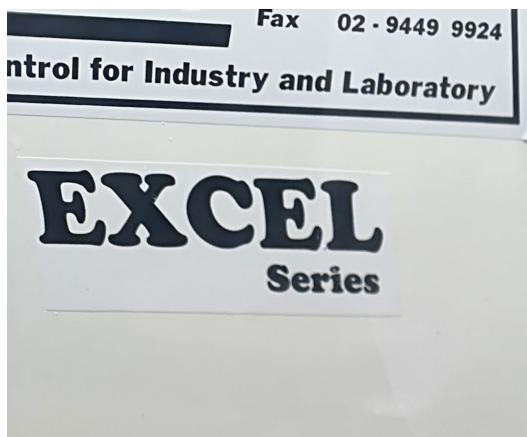
Steps followed:

```

```

For this experiment, I:

1. Prepared 10mMol NaCl, filtered using an air vacuum



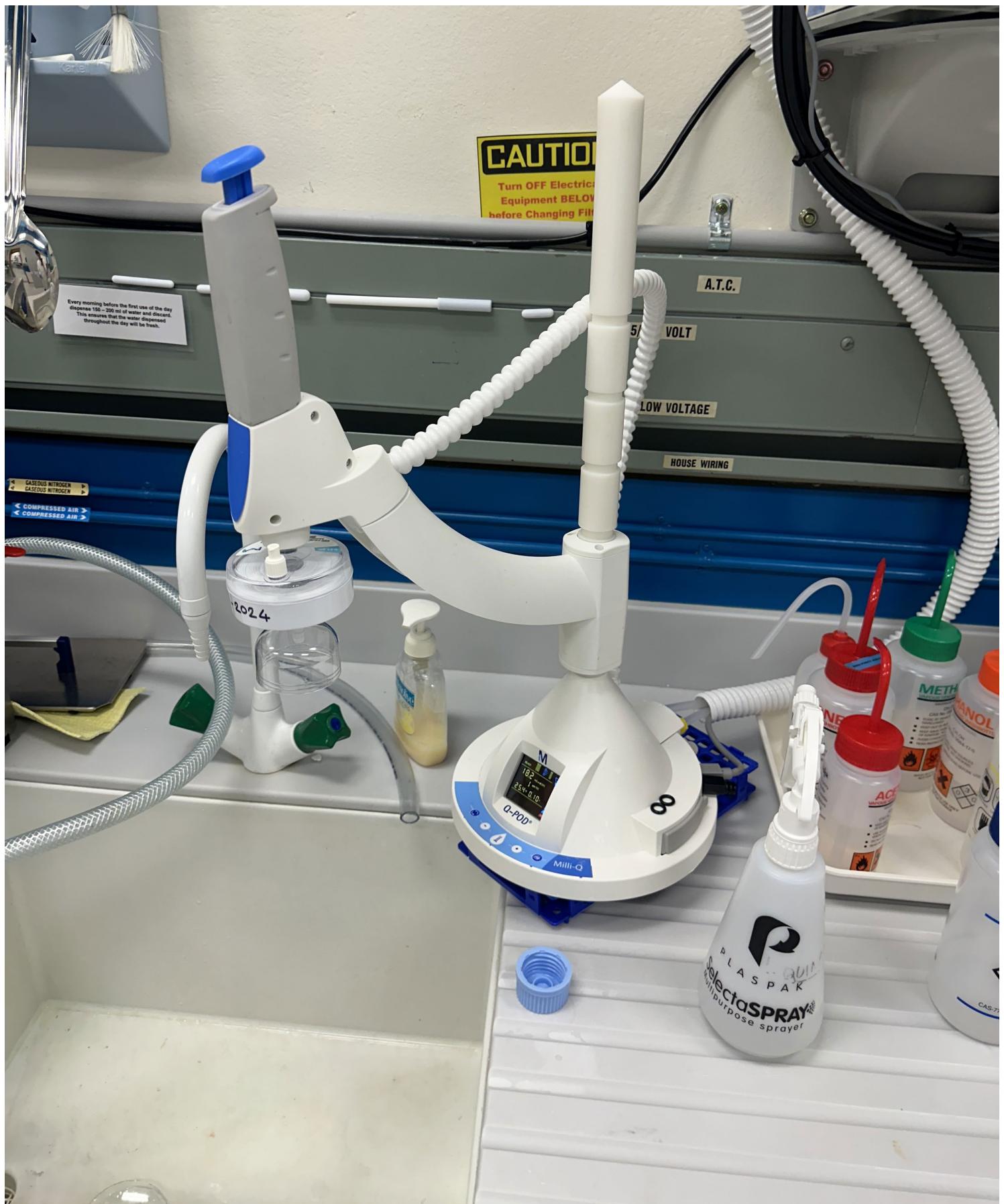




- The water used was from the MilliQ machine in the lab:







- The filters | used were:



Omnipore™  
0.1µm PTFE Membrane  
25 Filters, 47mm

Hydrophilic, Nonsterile

REF: JWP04700  
LOT:0000192923



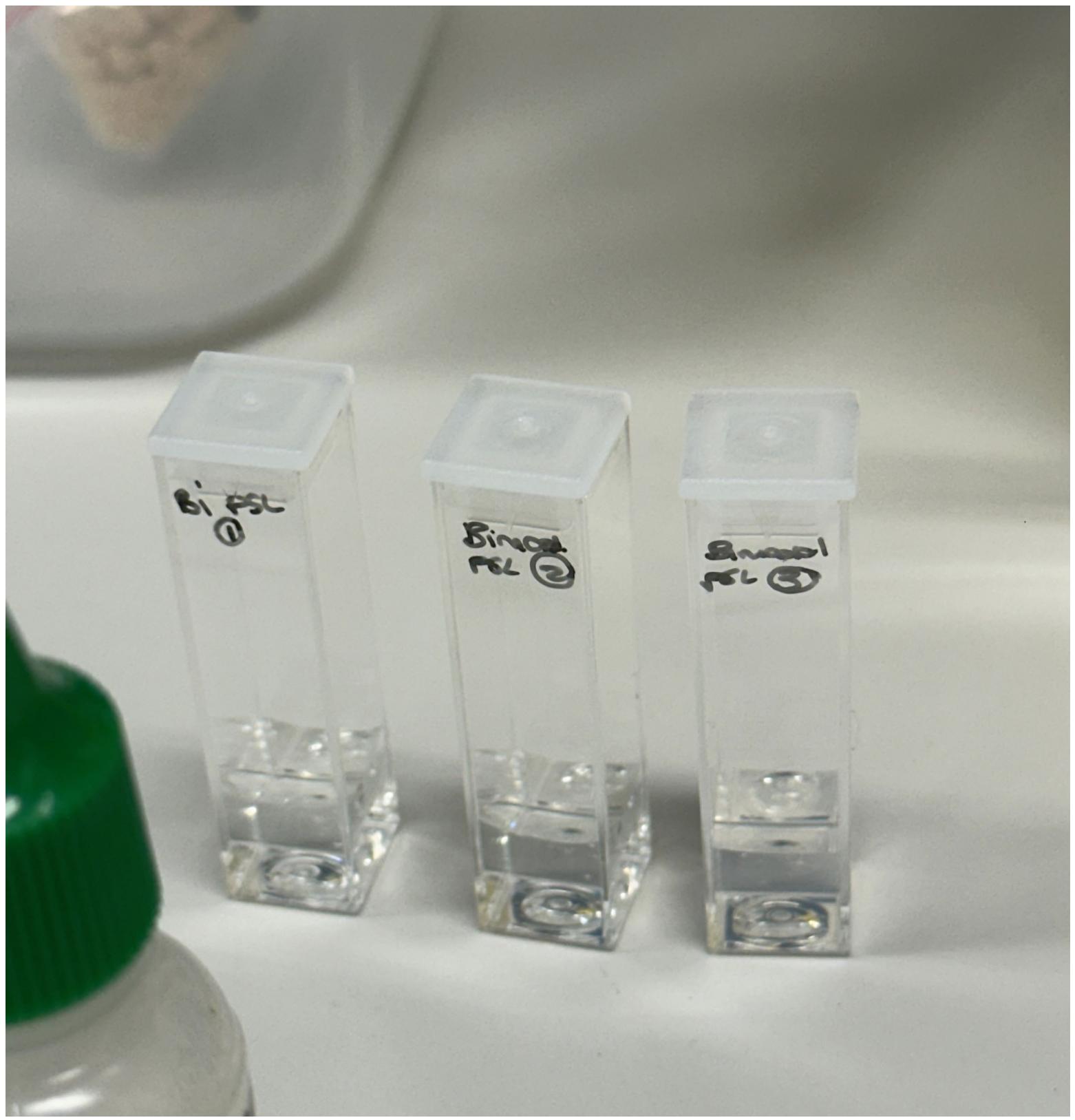
2. Pipetted a 20nm PSL into 4 cuvettes

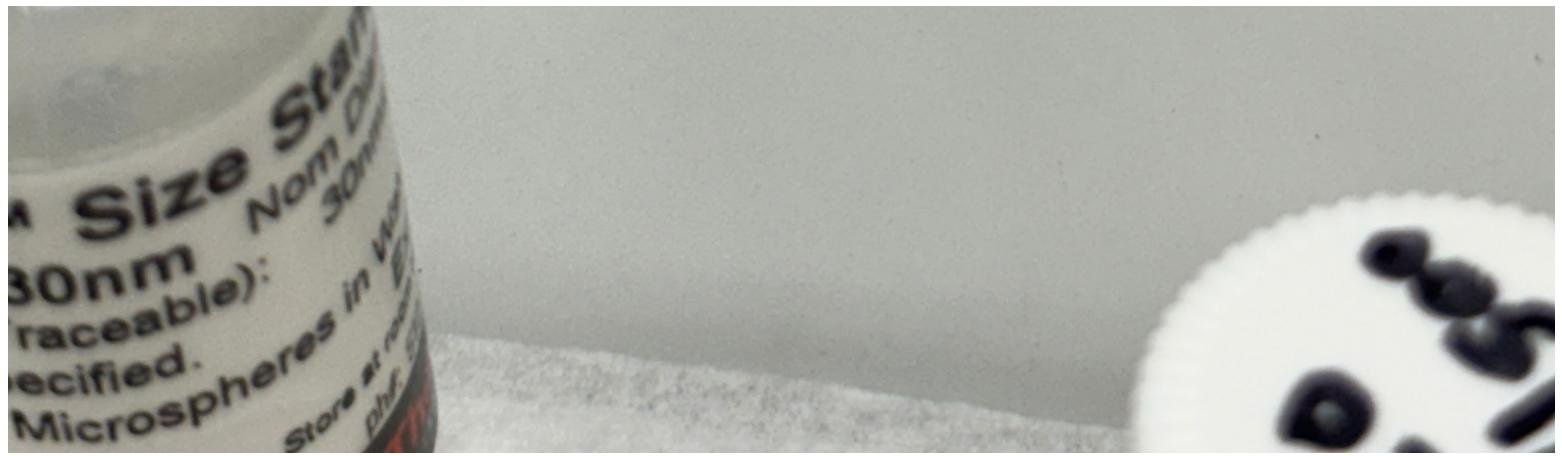




3. MD then explained the DLS machine and software

The cuvettes:





Notes from first run<sup>[10]</sup>

## Remainder of Wet Lab Work

Most of the remaining weeks was spent either writing code for the project, which was primarily done at home, or assisting MD in troubleshooting pressure build ups in the AF4 machine. All handwritten notes are uploaded to the repository, as well as photos of any experiments conducted.

## References and Annotated Bibliography

The list of references will continue to be expanded as I get more information.

Berne, B. J., & Pecora, R. (2000). Dynamic light scattering: with applications to chemistry, biology, and physics. Courier Corporation.

Brown, W. (Ed.). (1996). Light scattering: Principles and development. Clarendon Press.

Cho, T. J., & Hackley, V. A. (2010). Fractionation and characterization of gold nanoparticles in aqueous solution: asymmetric-flow field flow fractionation with MALS, DLS, and UV-Vis detection. *Analytical and Bioanalytical Chemistry*, 398(5), 2003–2018. <https://doi.org/10.1007/s00216-010-4133-6>

Guinier, A., Fournet, G., Walker, C. B., & Yudowitch, K. L. (1955). *Small-angle Scattering of X-rays*. New York: Wiley.

Kato, H., Nakamura, A., Takahashi, K., & Kinugasa, S. (2012). Accurate size and size-distribution determination of polystyrene latex nanoparticles in aqueous medium using dynamic light scattering and asymmetrical flow field flow fractionation with multi-angle light scattering. *Nanomaterials*, 2(1), 15–30. <https://doi.org/10.3390/nano2010015>

OriginLab. (2024). *Origin User Guide* (2024b). OriginLab Corporation.

Pecora, R. (2000). Dynamic light scattering measurement of nanometer particles in liquids. *Journal of Nanoparticle Research: An Interdisciplinary Forum for Nanoscale Science and Technology*, 2(2), 123–131. <https://doi.org/10.1023/A:1010067107182>

Podzimek, S. (2011). Light scattering, size exclusion chromatography and asymmetric flow field flow fractionation: powerful tools for the characterization of polymers, proteins and nanoparticles. John Wiley & Sons.

Quattrini, F., Berrecozo, G., Crecente-Campo, J., & Alonso, M. J. (2021). Asymmetric flow field-flow fractionation as a multifunctional technique for the characterization of polymeric nanocarriers. *Drug Delivery and Translational Research*, 11(2), 373–395. <https://doi.org/10.1007/s13346-021-00918-5>

StataCorp. (2023). *Stata 18 User's Guide* (18). Stata Press.

Terrill, N. (n.d.). Small Angle X-Ray Scattering. *Diamond Light Source*. <https://www.diamond.ac.uk/dam/jcr:fb6f93b8-2cd2-4646-b6ae-4a426d85c19b/Scattering%2520101.pdf&ved=2ahUKEwjQ0MvR64KJAxUtsVYBHaScNZ4QFnoECBUQAQ&usg=AOvVaw2q9wDc-eUOsCVIJrGjzlk>

Till, U., Gaucher-Delmas, M., Saint-Aguet, P., Hamon, G., Marty, J.-D., Chassenieux, C., Payré, B., Goudounèche, D., Mingotaud, A.-F., & Violleau, F. (2014). Asymmetrical flow field-flow fractionation with multi-angle light scattering and quasi-elastic light scattering for characterization of polymersomes: comparison with classical techniques. *Analytical and Bioanalytical Chemistry*, 406(30), 7841–7853. <https://doi.org/10.1007/s00216-014-7891-8>

Wyatt Technology. (n.d.). *TN6006: Eclipse and VISION Handbook*. Wyatt Technology Corporation.

Wyatt Technology. (2020a). *Astra User's Guide* (MM1111 Rev. A). Wyatt Technology Corporation.

Wyatt Technology. (2020b). *DAWN User's Guide* (M3220 Rev. B). Wyatt Technology Corporation.

## Footnotes

This is also a running list that I'm constantly updating.

1. Requires Refractive Index (RI) detector for most accurate measurement. Not available at NMI at the time of my placement. ↵
2. Outline in report what dilute means, if relevant. It is set-up dependent. ↵
3. State assumed value in report and why. ↵
4. Unsure if this one is used at NMI, I believe it is a default accessory. ↵
5. Report must include outline of all assumed / known values. ↵ ↵
6. Final report and Notebook are planned to include a method of measuring this once an RI detector is available at NMI. ↵
7. For final report and notebook, I need to double check what small and large means. ↵

8. Page 2 ↵
9. Lab Notes 1, page 4 of PDF ↵
10. Lab Notes 1, page 8 of PDF ↵