# Scorpy Progress Report

## Patrick Adams

# 1 Introduction

Scattering Correlation in Python (SCORPY) is a python package developed by members of the X-ray diffraction group in the department of Physics at RMIT University. The goal of the package is to supply easy to use tools for calculating and analysing scattering correlation data, with an emphasis for use with X-ray diffraction from protein crystals.

# 2 Definitions

## 2.1 Coordinates

Spherical coordinates are described using the ISO 80000-2:2019 convention. A point $p$ in spherical coordinates is given by $p = (r, \theta, \phi)$, where $r \geq 0$ is the radial distance from the origin, $\theta \in [0, \pi]$ is the polar angle, and $\phi \in [0, 2\pi)$ is the azimuth angle. All angles are measured in radians.

The polar angle $\theta$ is measured between a point and the positive $z$ axis (in rectilinear coordinates). For example, a point on the positive $z$ axis implies $\theta = 0$, a point on the negative $z$ axis implies $\theta = \pi$, and point where $z = 0$ implies $\theta = \frac{\pi}{2}$. Colloquially, theta is referred to as the "up-and-down" angle.

The azimuth angle $\phi$ is measured from a points projection onto the $(x, y)$ plane, and the positive $x$ axis. For example, a point that is perpendicular to the positive $x$ axis implies $\phi = 0$, a point that is perpendicular to the positive $y$ axis implies $\phi = \frac{\pi}{2}$, a point that is perpendicular to the negative $x$ axis implies $\phi = \pi$, and a point that is perpendicular to the negative $y$ axis implies $\phi = \frac{3\pi}{2}$.

The spherical coordinate system is illustrated in the diagram below.

Note that `pyshtools`, a python package for calculating spherical harmonic functions, uses a different standard of spherical coordinates. In `pyshtools`, angles are measured in degrees, and the polar angle $\theta$ is the angle of inclination above the $(x, y)$ plane. For example, a point on the positive $z$ axis implies $\theta = 90$, a point on the negative $z$ axis implies $\theta = -90$, and point where $z = 0$ implies $\theta = 0$.

## 2.2 Sampling

The sampling axis points within a 3D volume depends on if a given axis is periodic or not. We assume that for a given non-wrapped axis $x$, there is some minimum and maximum value

$x_{min}$ and $x_{max}$ that defines a continuous domain of the axis. This domain is then sampled at $n_x$ discrete points along the axis, across a series of $n_x$ bins. The sample point is defined at the centre of each bin, where the bin width $\Delta x$ is calculated from the following equation;

$$\Delta x = \frac{|x_{max} - x_{min}|}{n_x}$$

For example, if a non-wrapped axis domain is between 0 and 1, and there are 10 bins or sample points across the domain, then the first bin covers a subdomain $[0, 0.1)$ and the sample point is 0.05. Any calculated value within this subdomain is "placed" within the bin. The next bin covers a subdomain $[0.1, 0.2)$, with a sample point of 0.15. This continues until the last bin, which has an extended subdomain to include the maximum value of the axis. The subdomain of the last bin will be $[0.9, 1]$ with a sample point of 0.95.

For wrapped axis, we require that the $x_{min}$ and $x_{max}$ values are equivalent. In this case, the bin width $\Delta x$ is calculated as above. However, the first sample point is at $x_{min}$, where the first bin covers a subdomain from $\left[-\frac{\Delta x}{2}, \frac{\Delta x}{2}\right)$.

A diagram of explaining the sampling modes is shown below.

## 2.3 Correlations

Let $W$ be a set of 2D polar diffraction patterns $I_\omega(q, \phi)$ of identical single particles in random orientations, where $\omega$ describes orientation of the particle. If we assume that the set $W$ is large enough to span every orientation of the particle, then the scattering correlation function $C(q_1, q_2, \Delta\Psi)$ is given by the equation;

$$C(q_1, q_2, \Psi) = \sum_{\omega=0}^{W} \int_0^{2\pi} I_\omega(q_1, \phi) I_\omega(q_2, \phi + \Psi) d\phi$$

## 2.4 Calculating Correlations

# 3 Object Overview

## 3.1 Readers

Reader objects handle the input data for calculating correlations.

### 3.1.1 CifData

### 3.1.2 PeakData and ExpGeom

## 3.2 Vols

In scattering correlation analysis, 3D data sets frequently occur in the form of PADFs ($C(r_1, r_2, \Psi)$), Correlation functions ($C(q_1, q_2, \Psi)$) and Harmonic Correlation functions ($B_l(q_1, q_2)$). Vol objects are a representation of any 3D data set or space.

**3.2.1**

# 4 Progress

## 4.1 Verifying Correlations

Correlation volumes can be calculated with a variety of methods, depending on the source of data. We will focus on the calculation

## 4.2 Calculating CIF Correlations by Hand