Data Formats Working Group

canSAS XIV – Taipei, November 3rd 2024

Overview & Members

About the group

A common data format has been at the heart of canSAS discussions since the formation of the collaboration in 1998.

This working group was formed at the cansas-V meeting at NIST in 2005 and a first canSAS format for 1d I vs Q data was defined using XML.

In 2012 a working meeting (”hackathon”) was held in Uppsala and a proposal for what would become NXcanSAS (NeXus supported, using HDF) was developed.

NXcanSAS was ratified by the NeXus International Advisory Committee in October 2016 and v1.0 is the current operational standard (https://manual.nexusformat.org/classes/applications/NXcanSAS.html)

Members

- Andrew Jackson (ESS)- Chair/Convenor
- Jeff Krzywon (NIST)
- Pete Jemian (APS)
- Steve King (ISIS)
- Andrew Nelson (ANSTO)
- Jan Ilavsky (APS)
- Lisa Debeer-Schmitt (ORNL)

Remit

Development of common reduced data format for SAS and SAS related techniques:

- Support for the NXcanSAS data format, including release of new format versions and interface with the NeXus committees
- Supporting extensions of NXcanSAS
- Supporting the development and implementation of formats for other methods as needed by the canSAS community

Magnetic/Polarised SANS

The Magnetic-SANS Problem

To reliably and unambiguously interpret Magnetic-SANS data it is necessary to know the spin history of the neutrons during the measurement. It makes sense to store this in the reduced output file as metadata which, in turn, requires NXcanSAS to:

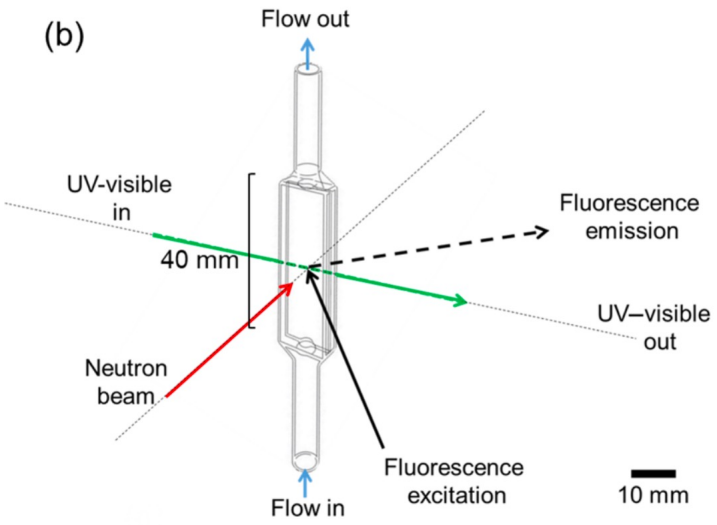
- have a way to include the polarization (aka spin state) information with the SANS data;
- have a way to denote the polarization vector as this defines the geometry and physical interpretation of the scattering cross section;
- have a way to identify the type of Magnetic-SANS technique that generated the data, and;
- to include metadata relating to the spin manipulating devices (order encountered, type, etc) used during the measurement of the data.

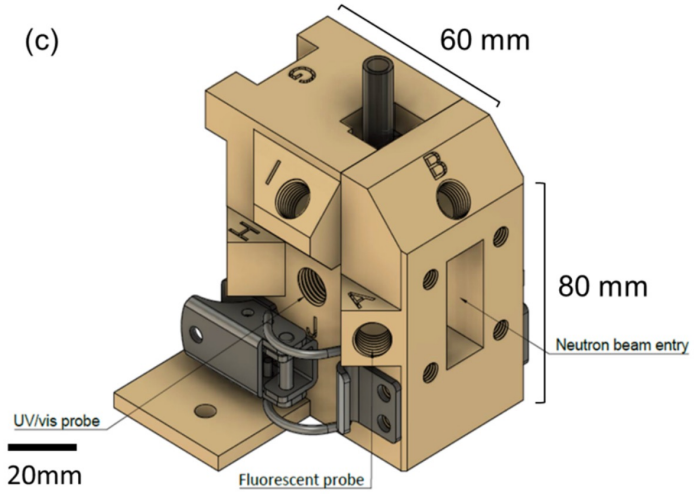
NXcanSAS v1.0 and NeXus provide for some, but alas not all, of this information.

Proposal : https://wiki.cansas.org/index.php?title=NXcanSAS_v1.1

Complementary Data

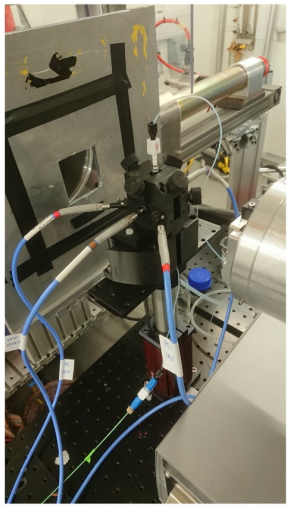
- Work started to support the NuRF setup (Dicko et al Rev. Sci. Inst. 91 (2020) https://doi.org/10.1063/5.0011325)at ESS as part of EU funded project SRESS3.
- Simultaneous SANS, UV-Vis and Fluorescence
- Used at ISIS and ILL for initial experiments

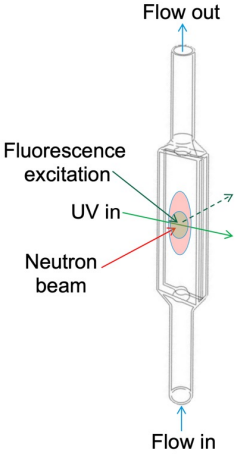




Experimental Set-up at Larmor (ISIS)

1. Knauer HPLC pumps for **sample mixing**
2. Flow cell with Ocean Optics **fluorescence** and **absorption** fibres optics connected
3. Anton Paar **density meter**





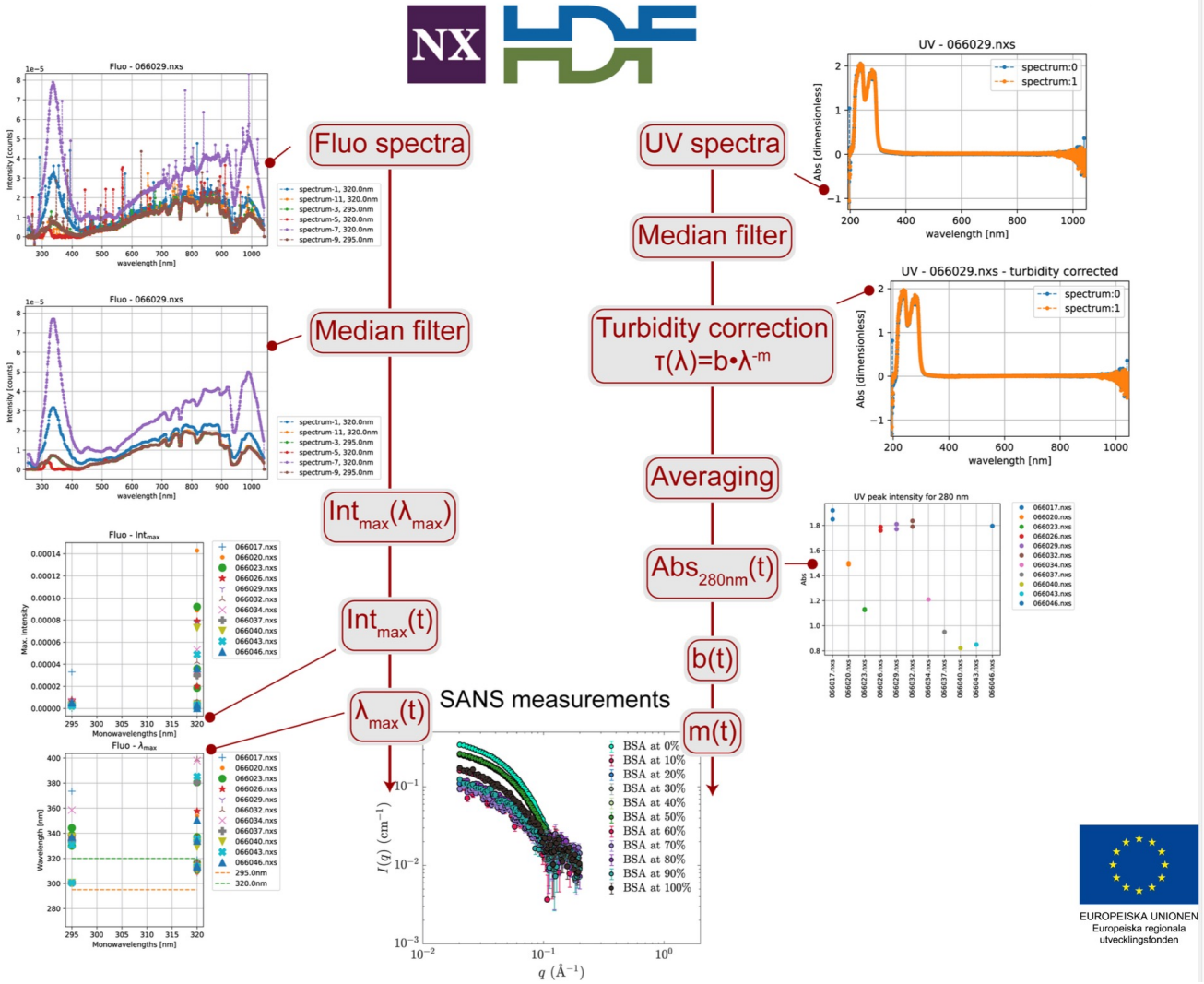
Complementary Data

Project Objectives :

- Dedicated python scripts
- Data format to capture multi-dimensionality
- NeXus compliance
- Scipp integration
- Modularity

Project Outcomes :

- Data format for auxiliary data has been established
- Proof of principle for data storage, reading, and processing/analysis using ESS scipp workflows
- Proof of principle experiments performed (BSA + warfarin)



To Do / Open Questions

- Proposal for magnetic/polarised to be ratified and submitted to NIAC – initial discussions started
- Support for complementary technique data (e.g. in-situ spectroscopies) – initial work done, need to develop proposal for incorporation into NXcanSAS
- Support for more advanced resolution definitions (see TOF WG presentation) – initial work done, needs more development and then proposal for incorporation into NXcanSAS (necessary for ESS)
- How (if?) to incorporate other SAS and related techniques – extension of NXcanSAS? or separate format?:
  - SESANS (working group exists)
  - Grating Interferometry
  - GISANS (working group exists - working with ORSO?)
  - XPCS?
  - Soft x-rays (RSoXS)?

Possible “hackathon”/WG meeting in 2025 to work on these topics – Planning for ESS to host, but other offers welcome!

Backup/Discussion

SESANS data …

From Greg Smith (ISIS)

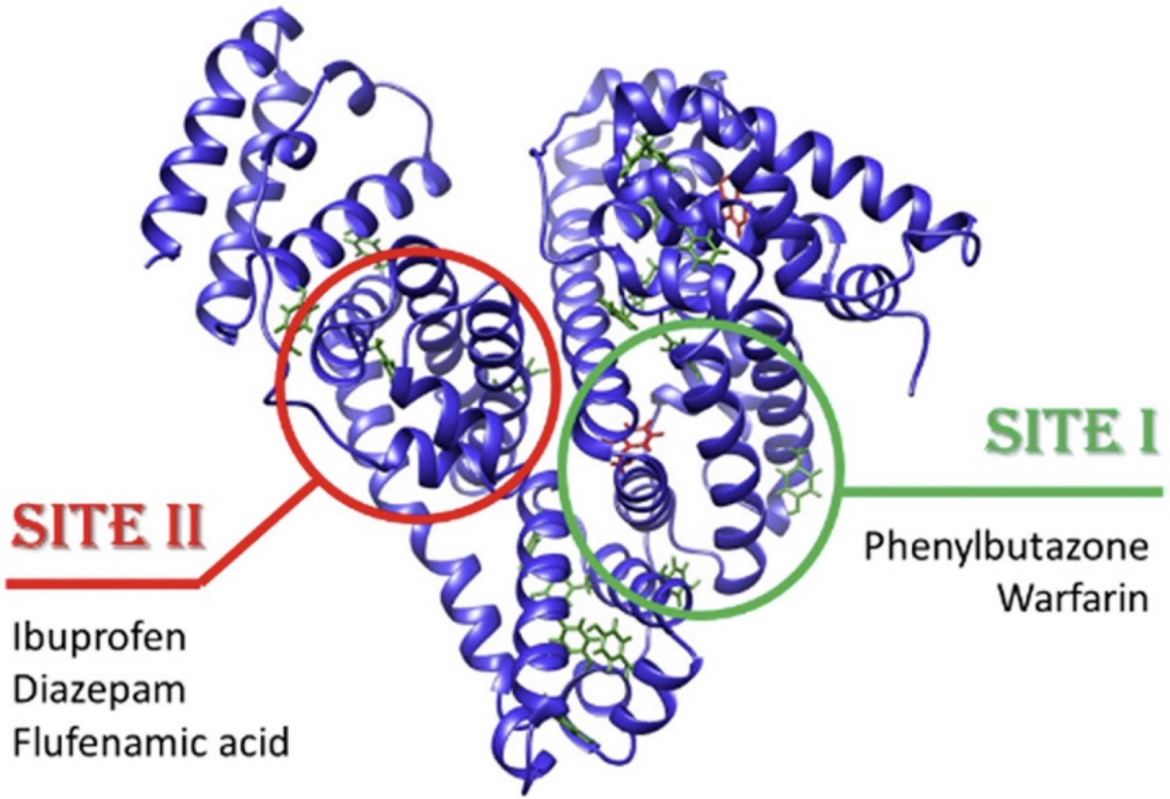
- What do we want the SESANSdata block to be? The rawest data that we measure is P/P0 (sample polarisation over blank polarisation), but to make that instrument independent, we normalise it by wavelength and thickness. In the current text-based format, we do this before the file is written (so it's the normalised scattering correlation function, not P/P0), but if you have all the metadata, then you could this when importing to fitting. However, I'd like to avoid users taking a P/P0 block for plotting without ensuring it's normalised.
- What things would we like to use to correct our data? Given finite detector sizes and conventional SANS, we already sometimes need to correct for scattering missing the monitors (transmission monitor and downline detector). Are their other metadata that we need to record to ensure that these corrections can take place? These metadata will likely be different than conventional SAS/VSAS/USAS and may not already be defined by the format.

NURF applied to BSA

UV absorbance to monitor **BSA concentration, warfarin insertion and turbidity**

Fluorescence emission to monitor **BSA warfarin binding and changes in the BSA tertiary structure** (denaturation)

SANS to monitor the **shape, internal organization**



W. Potrezebowski, J. Houston, A. Rennie, C. Dicko

Integration into scipp

Demonstration of NUrF function module

The Nurf Python module contains functions to extract UV and fluo data from a LoKI.nxs file.
It is capable of correcting the measured spectra for reference and dark current. It calculates the final fluo and UV spectra and p
them. A median filter can be applied to the spectra.
For UV the following parameters can be extracted:

- Absorbance at 280nm
- Turbidity and m, b factor

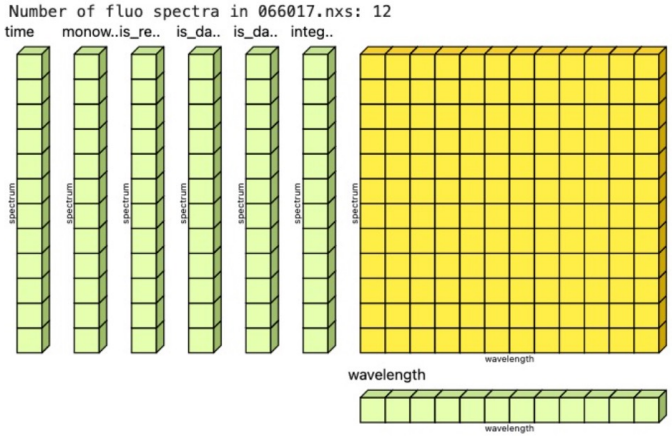
For fluo the following parameters can be extracted:

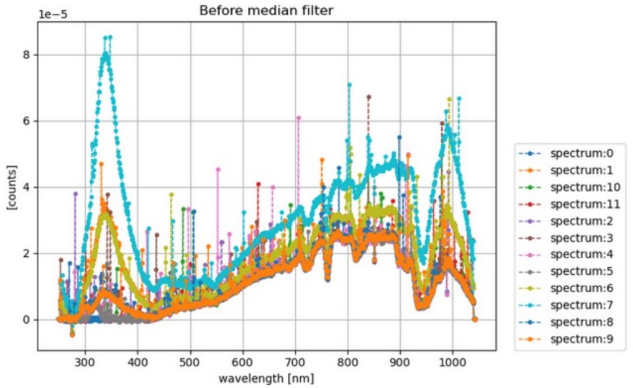
- Maximum peak intensity
- Peak position corresponding to the maximum peak intensity

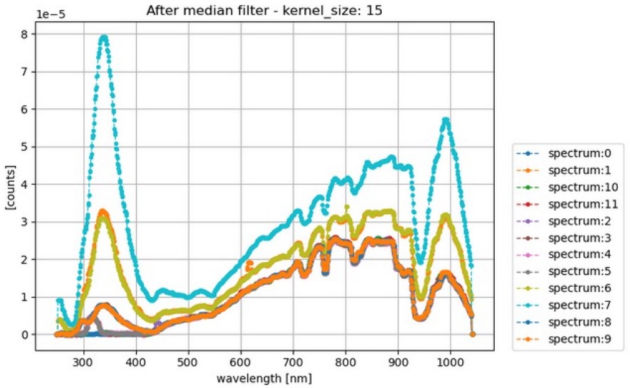
Besides standard Python libraries like matplotlib, os, scipy, and numpy, it is relies on scipp, scippnexus, and scippneutron.

Prepare the stage and load the required tools.

```
from nurf import *
from ill_auxilliary_funcs import *
from scipp.signal import butter, sosfiltfilt
```







W. Potrezebowski, J. Houston, A. Rennie, C. Dicko

EUROPEISKA UNIONEN
Europeiska regionala utvecklingsfonden

Instrument control integration

It is complicated…