

# Reference-Quality Data Corrections with MoDaCor

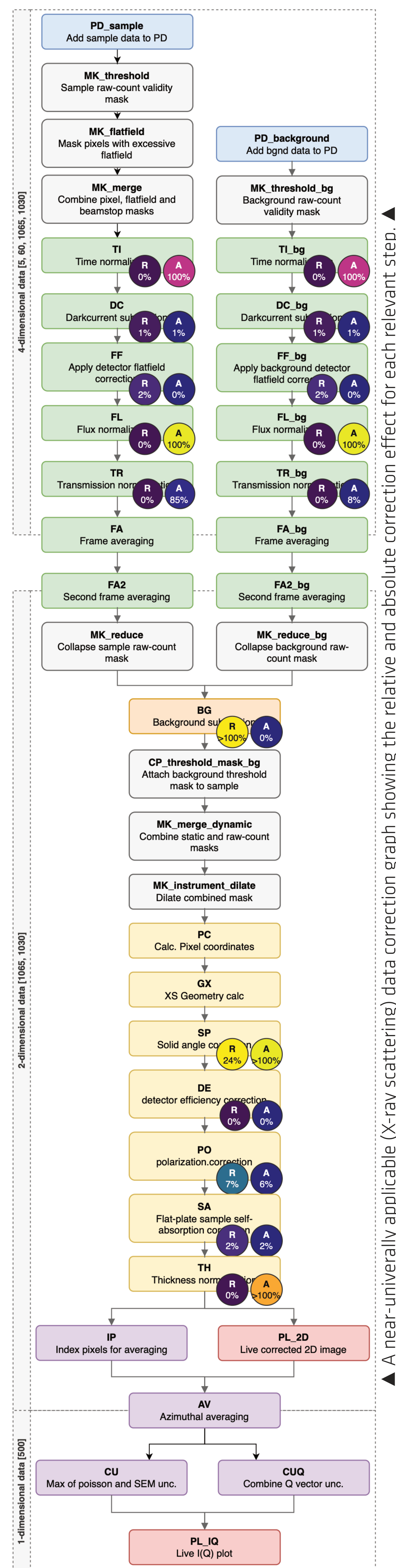


Bundesanstalt für  
Materialforschung  
und -prüfung

Brian R. Pauw<sup>1</sup>, I. Breßler<sup>1</sup>, A.F. Hörmann<sup>1</sup>, J. Kieffer<sup>2</sup>, A. Moser<sup>3</sup>, G. J. Smales<sup>4</sup>, T. Snow<sup>5</sup>, M. Storm<sup>6</sup>

1) Bundesanstalt für Materialforschung und -Prüfung (BAM), Germany  
2) European Synchrotron Radiation Facility (ESRF), France  
3) Anton Paar GmbH, Austria

4) Technical University of Graz (TU Graz), Austria  
5) Diamond Light Source, UK  
6) Helmholtz-Zentrum Hereon, Germany



## Benefits:

The quality of an experiment can only be guaranteed if *all* of its parts are consistently performed to a high standard. For the data corrections, MoDaCor therefore implements:

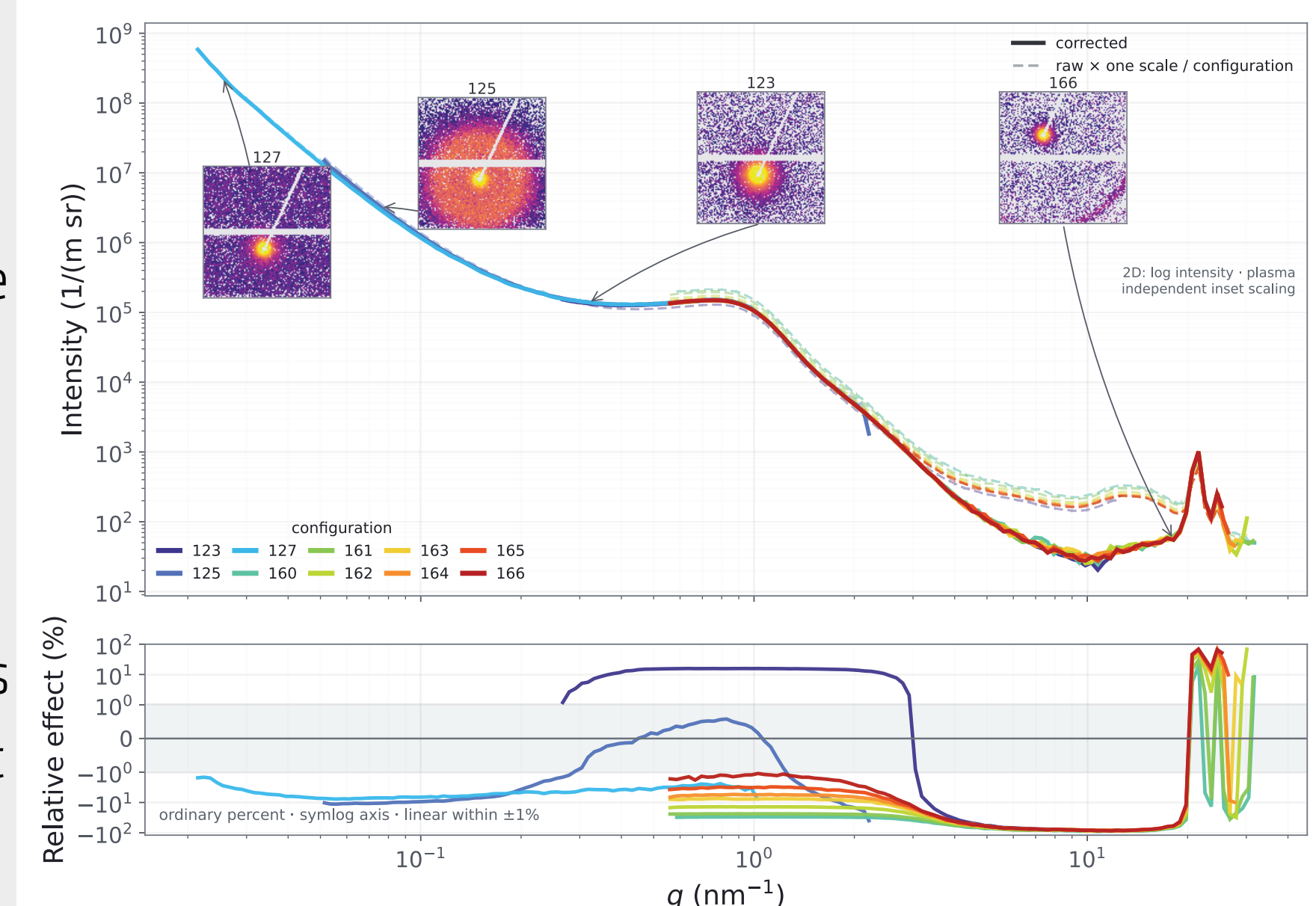
- Extensible, graph-driven modular corrections
- Units handling and validation
- Propagation of (multiple) uncertainty estimates
- Full traceability of the multidimensional corrections
- Server-based Web API with Tiled and HDF5 support
- Extensive, LLM-compatible documentation

## Applications:

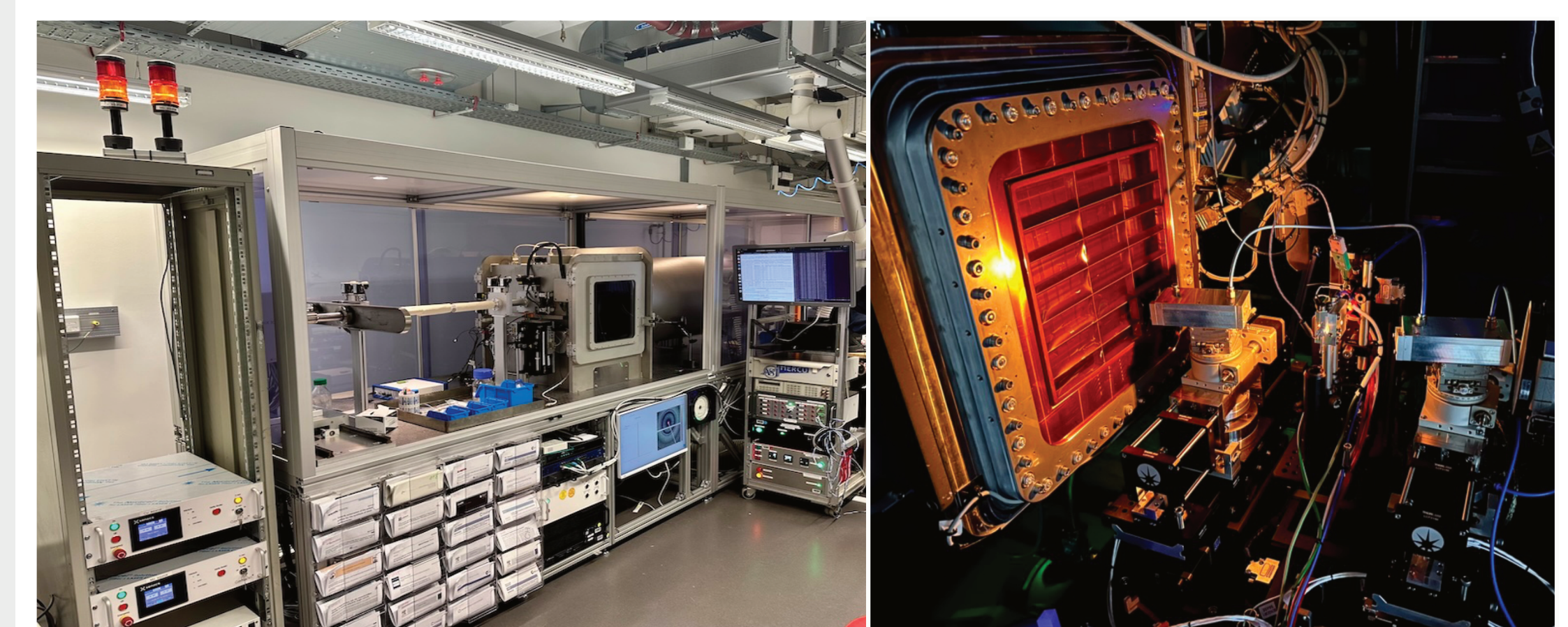
MoDaCor can be used to correct any non-time-of-flight data from the fields of scattering, diffraction and imaging. The modular corrections can be composed to form correction pipelines applicable to a near-universal range of samples.

Current application examples include transmission- and grazing-incidence X-ray scattering pipelines for laboratory- as well as synchrotron instruments, both for whole-dataset as well as chunked processing.

New application cases can be rapidly (interactively) deployed using the MoDaCor-example library, representative measurements and a coding LLM.



MoDaCor corrects multidimensional data to absolute units, here shown on data from a strongly-scattering sample. The data has been collected over a range of instrument configurations. Insets show corresponding corrected 2D data



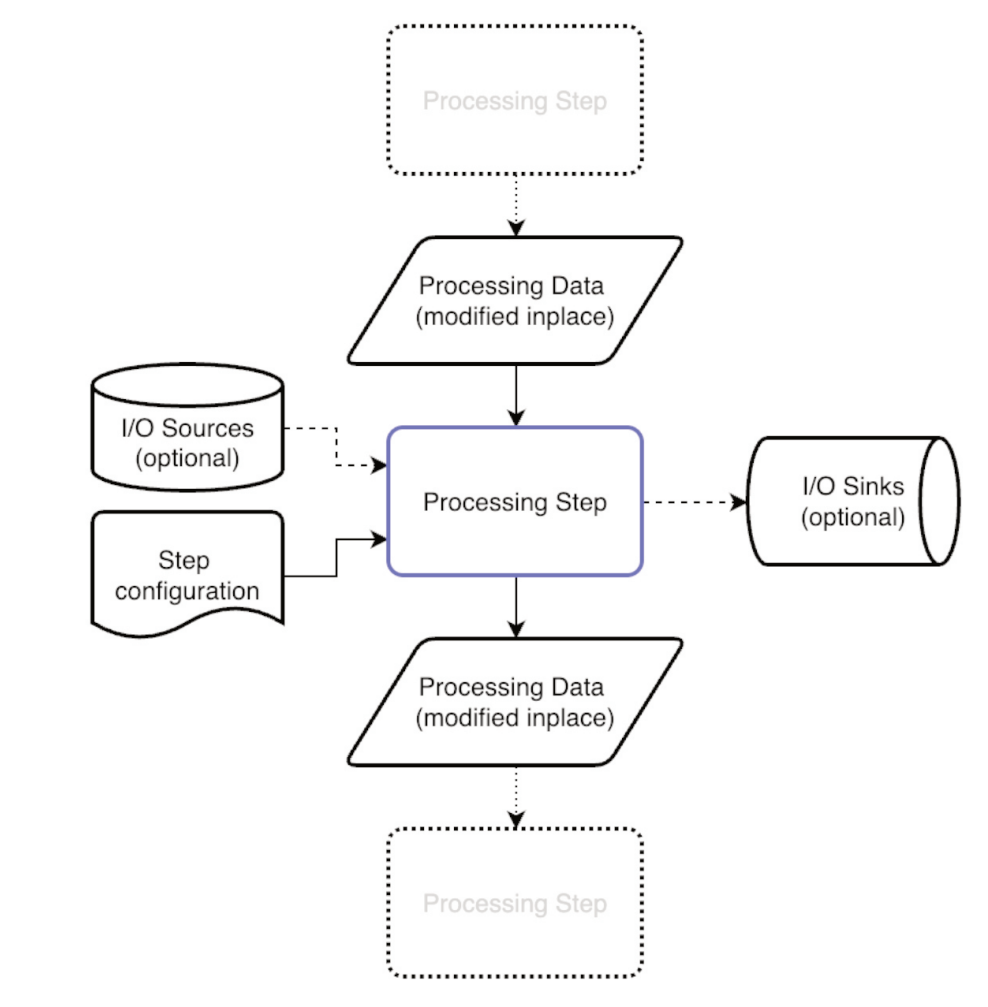
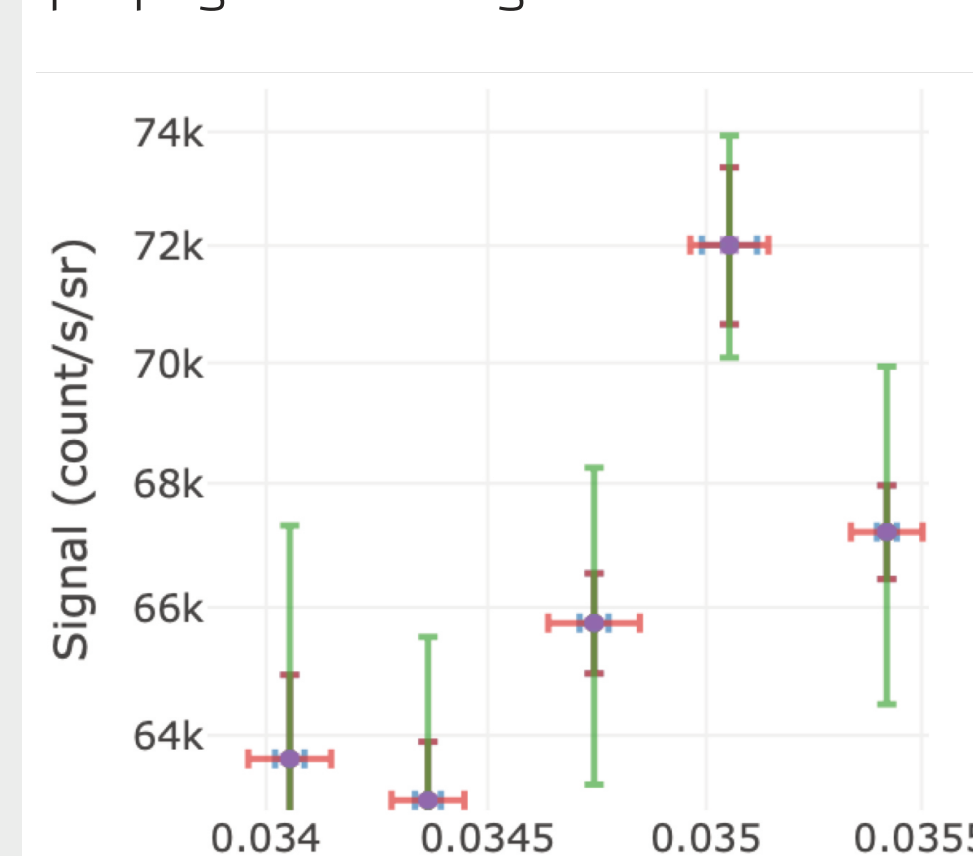
Useful for ▲ laboratory- as well as synchrotron ▲ experiments.

## Corrections:

A wide range of self-documenting generic and scattering-related correction modules have been implemented already. Highlights include:

- A module that traverses NeXus transformations to calculate detector pixel coordinates in 3D in the lab frame of reference
- An advanced background- and absorption correction module for (flow-through) capillary-contained samples, using reverse ray-tracing.
- A generic masking module that can mask detector pixels based on any other detector-sized datasets, such as flatfields, geometry or uncertainty maps.

Multiple uncertainty estimates can be propagated during the corrections ▼



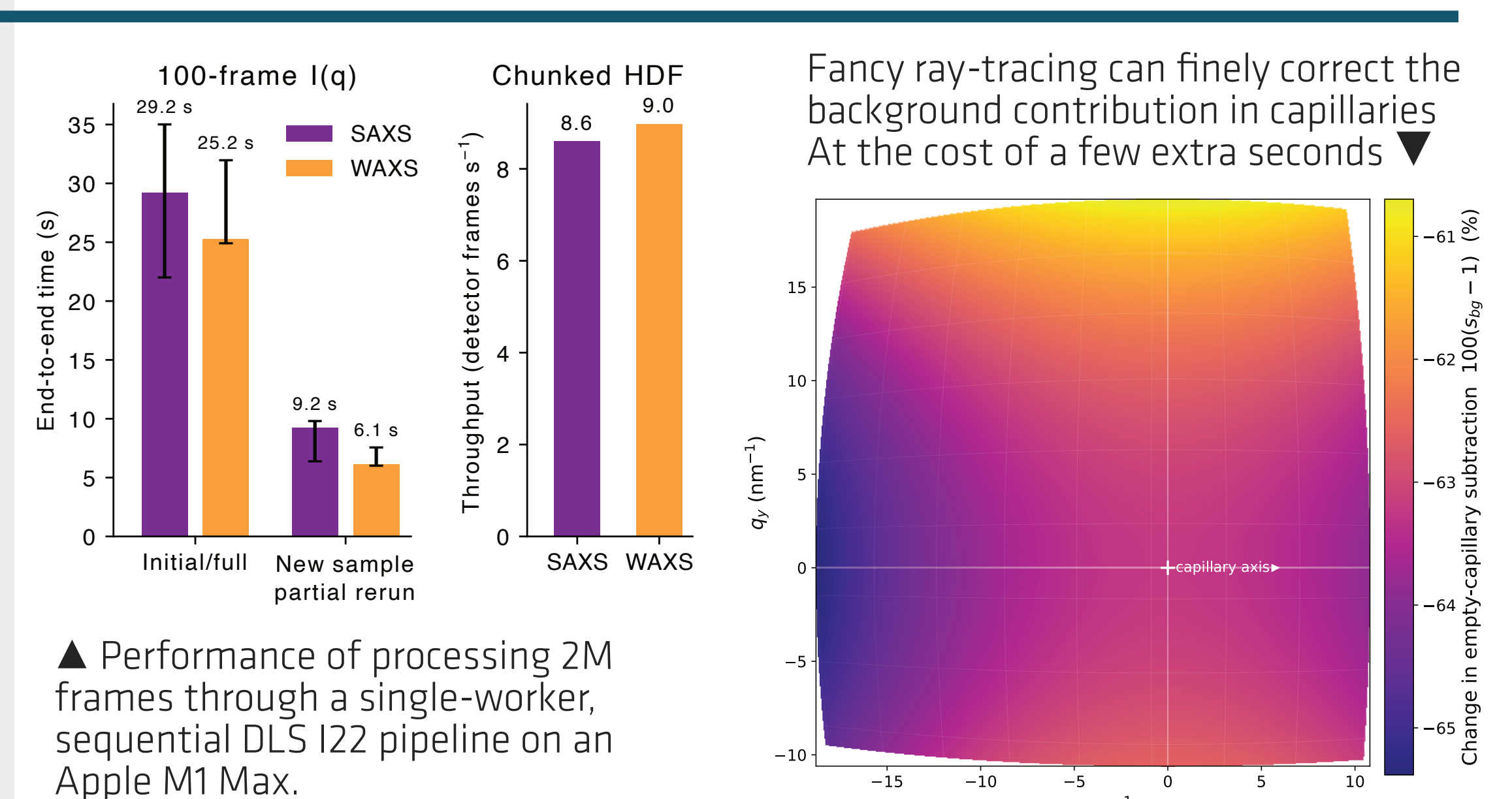
▲ Processing steps have a universal layout: Processing data is modified inplace (for performance) and all have access to I/O.

## Performance:

Initial testing shows MoDaCor to be ~2-3x faster than a corresponding pipeline in the Data Analysis WorkbeNch (DAWN), despite additional work on validation and propagation per step.

This is due to the partial graph rerun of graphs for new data; unchanged data is not recalculated. This particularly helps with clever pipeline designs.

If much faster performance is needed, MoDaCor can be used as a reference benchmark; for a subset of measurements, it can validate faster, more bespoke data correction implementations (such as PyFAI) within the actual application envelope.



▲ Performance of processing 2M frames through a single-worker, sequential DLS I22 pipeline on an Apple M1 Max.

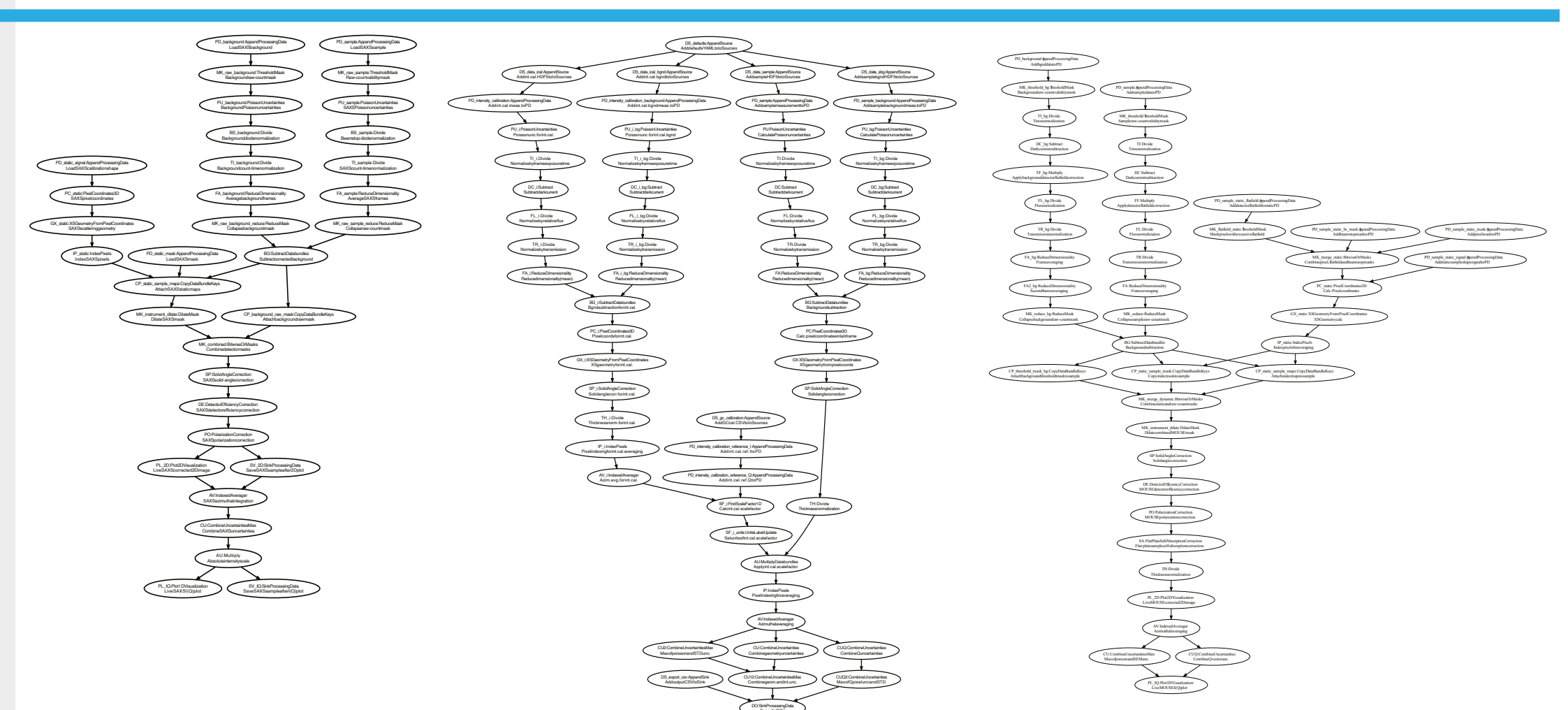
## Future work:

We will gradually increase the library of application examples to further demonstrate the versatility and applicability of the framework, while extending the correction module library.

Additional personnel resources permitting, a user interface leveraging the already existing graph I/O tools will be made to facilitate pipeline modification and testing.

Closer integration with workflows using Bluesky and Tiled will allow automated data corrections with a minimum of extra plumbing.

Join in on the fun today!



▲ Each application may require different graph topologies. From left to right: DLS I22 SAXS pipeline, BAM SAXSess reference pipeline, BAM MOUSE operando pipeline. Far left on the poster: MOUSE absolute intensity pipeline

Sicherheit in Technik und Chemie