ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-TUPAB413

RAPID BROWSER-BASED VISUALIZATION OF LARGE NEUTRON SCATTERING DATASETS*

K. C. Bruhwiler†, D. L. Bruhwiler, R. Nagler, P. Moeller, RadiaSoft LLC, Boulder, CO, USA
C. Hoffmann, A. T. Savici, Z. Morgan, M. Tucker, Neutron Sciences, ORNL, Oak Ridge, TN, USA
A. Kuhn, J. Mensmann, P. Messmer, M. Nienhaus,
S. Roemer, D. Tatulea, NVIDIA, Berlin, Germany

Abstract

Neutron scattering makes invaluable contributions to the physical, chemical, and nanostructured materials sciences. Single crystal diffraction experiments collect volumetric scattering data sets representing the internal structure relations by combining datasets of many individual settings at different orientations, times and sample environment conditions. In particular, we consider data from the single-crystal diffraction experiments at ORNL. A new technical approach for rapid, interactive visualization of remote neutron data is being explored. The NVIDIA IndeX 3D volumetric visualization framework is being used via the HTML5 client viewer from NVIDIA, the ParaView plugin, and new Jupyter notebooks, which will be released to the community with an open source license.

NEUTRON SCATTERING DATA

The data in this paper represent neutron scattering from a single crystal neutron diffraction instrument. The object of the study is to investigate the material structure a 3-dimensional periodic arrangement of atoms and elements in a highly ordered single crystal sample. The scattering pattern of this array, which is detected with scintillator based detectors, is related to the original structure through a Fourier transformation. This allows calculations of the atomic positions [1], to solve the structure – property relations in basic and applied material science [2]. A typical data set for visualization is 500 MB to 100 GB in size. For complete data visualization and analysis, a series of individual data sets at many sample orientations are combined, to provide a 360 degree representation of the sample structure.

TOPAZ

The TOPAZ [3] instrument (see Fig. 1) is located at the Spallation Neutron Source (SNS) at Oak Ridge National Laboratory (ORNL). The neutrons for scattering experiments are produced by spallation: a proton beam of approximately 1.4 MW power hits a mercury target 60 times per second, strips excess neutrons, which are then moderated to thermal energies. This provide neutrons with favorable kinetic energy ranges for scattering interactions (~ 10 meV – 1000 meV). Thermal neutrons are channeled through a guide to the prepared sample inside a detector tank. Scattered neutrons are detected in a time resolved mode at arrival, according to the kinetic energy of each neutron in "neutron event" mode. The collected neutron events make up a data set of scattering signal from the sample and scattering from other sources of background. The signal has a variation of periodic signatures at multiple levels of statistical significance to be extracted and visualized for analysis [4].

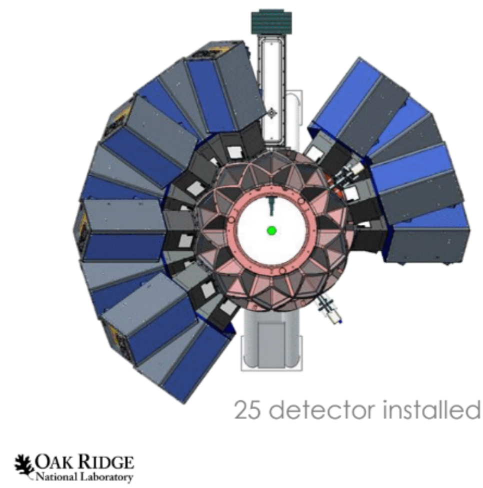

Figure 1: The TOPAZ instrument.

MANTID

Mantid algorithms are used to process single crystal diffraction data at SNS [5]. Time of flight data collected from TOPAZ instrument is converted to a multi-dimensional events workspace, by transforming the measured data (scan) into reciprocal volume with coordinate system attached to innermost axes of the instrument goniometer. Multiple scans can be merged into a single dataset, to map the full reciprocal space volume [6]. The data is then converted into a dense representation, a weighted histogram of equally spaced bins. The weights for normalization are calculated using a standard scatterer (vanadium). The data used in this work focuses on scans from a variety of experiments ranging from standard samples used for calibration, to Benzil. Benzil has a nonorthogonal crystal lattice and exhibits diffuse scattering.

The Mantid program contains a variety of tools for preprocessing data and effectively visualizing 1D and 2D histogrammed slices, however they are deficient in visualizing 3D datasets. Previously, Mantid integrated ParaView [7] to visualize neutron scattering datasets in 3D, however the sys-

* This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0021551.
† kevincb@radiasoft.net

tem was not user friendly and proved too slow for effective analysis, and was subsequently removed.

NVIDIA INDEX

NVIDIA IndeX [8] is a tool for interactively visualizing large volumetric datasets. IndeX utilizes a cluster of (optionally distributed) GPUs to render its visualizations, allowing it to visualize an effectively unlimited amount of data. With IndeX, users are able to dynamically edit the color and transparency of the data based on the values of the volumes, as well as zoom to regions of interest and occlude portions of individual datasets. IndeX's ability to visualize large voxel datasets (datasets which are organized into an equally spaced 3D grid) makes it ideal for visualizing neutron scattering datasets in which data is already required to be binned.

ParaView Integration

IndeX also has a ParaView supported plugin [9] which supports both local and remote viewing. Using IndeX through ParaView gives users access to all the features ParaView provides, including tools for slicing and rotating datasets. However, ParaView integrations comes with two major drawbacks. First, the ParaView interface is complex and difficult to learn, making it potentially inaccessible to researchers who lack programming experience. Second, adding support for features required for neutron scattering analysis is difficult, as it must interface with both ParaView and the IndeX plugin. Consequently, our efforts both in this and in future work will focus on developing an IndeX application tailored to the needs of neutron scatter analysis.

VISUALIZATION WORKFLOW

The visualization workflow for TOPAZ data can be broken-up into three distinct phases: (1) feature identification, where the raw data is processed to determine the periodicity of the scattering pattern is determined, which can be used to intelligently select sub-regions or slices of data to view, (2) pre-processing, in which the raw data is converted into a voxel format which can be interpreted by IndeX and stored, and (3) visualization, where the data is interactively shown.

Feature Identification

One of the primary features of interest in neutron scattering data is Bragg peaks [4], regions which contain a much higher density of time-of-flight events than the surrounding area. The periodicity of these peaks along various axes is the key factor in determining the structure of the sample material. The periodicity can be determined by drawing lines that pass through multiple peaks and computing the minimum factor of the distances between points along a line. Additionally, these lines enable a transformation of the dataset to a space in which the peaks lie along the cardinal axes, which allows more intuitive visualization and simplifies future analysis. The result of this process can be seen in Fig. 2, alongside IndeX's Jupyter integration.

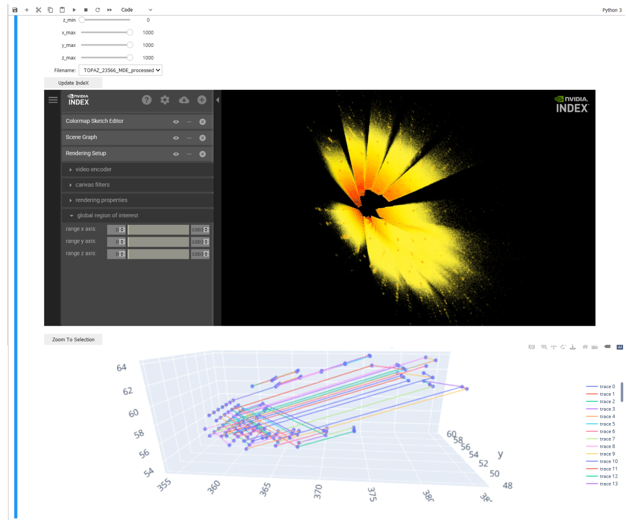

Figure 2: IndeX with Jupyter Notebook Integration. Shown with a custom UI alongside a seperate user analysis.

Data Pre-Processing

Before being rendered in IndeX, TOPAZ data must first be converted into voxel format. This is done by normalizing the neutron time-of-flight events to a given range, based on the desired resolution of the visualization, rounding the event locations to the nearest whole value, and counting the number of events that fall into each cell of a 3D grid. This 3D grid can be stored as a matrix and loaded dynamically at user request. Cells, or groups of neighboring cells, with a large number of events (and consequently a high total value) will stand out clearly in the visualization, while cells with low values can be set to be transparent.

The primary challenge with using voxels as a data storage method is that they are very inefficient with sparse datasets. For example, in the Benzil sample TOPAZ data stored in a 1000^3 voxel grid more than 99.9% of the cells contain no time-of-flight events at all. Storage space can be drastically reduced by using sparse voxel datasets, in which only cells than contain values are stored and cells about which there is no information are assumed to be empty. IndeX supports Open Voxel Data Base (OpenVDB) [10], which uses a sparse voxel format.

Data Storage Benchmarks The Benzil TOPAZ sample consists of 17 rotations, each 15 GB in size making it 255 GB in total. Figure 3 compares the amount of disk and RAM required to store and load the voxel datasets. The size of dense voxel datasets is invariant, as every cell requires the same amount of space to store regardless of value, while sparse voxel datasets vary based on the number of non-empty cells. In Fig. 3 we show the size of the sparse voxels with all rotations included. A sparse voxel dataset with only one of the Benzil rotations would be $\sim 1/17^{th}$ the size. It can be seen that the dense voxel dataset reduces the total datasize to 0.005% of its original size, while the sparse voxel dataset reduces it to less than 0.00004% of the original size.

12th Int. Particle Acc. Conf. IPAC2021, Campinas, SP, Brazil JACoW Publishing
ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-TUPAB413

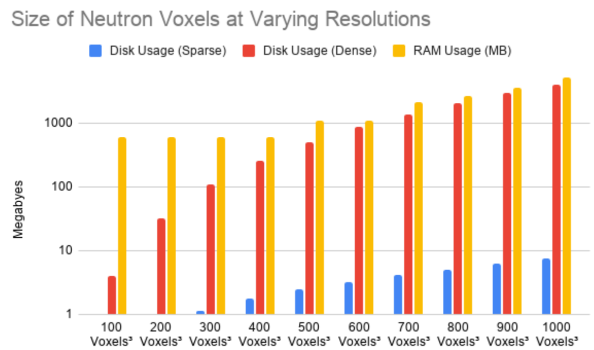

Figure 3: The storage space required by voxel datasets at varying resolutions to store the TOPAZ Benzil sample data. Note that the vertical axis is in log scale.

Visualization

IndeX visualizes voxel datasets interactively, allowing users to dynamically manipulate the color and transparency of the visualization based on voxel value, as well as change the perspective. It does this by performing the rendering operations on a remote server and streaming a video back to the user. User interactions are sent to the server, which adjusts its operations accordingly. Figure 4 shows how colors and transparency can be adjusted to better view features such as Bragg peaks.

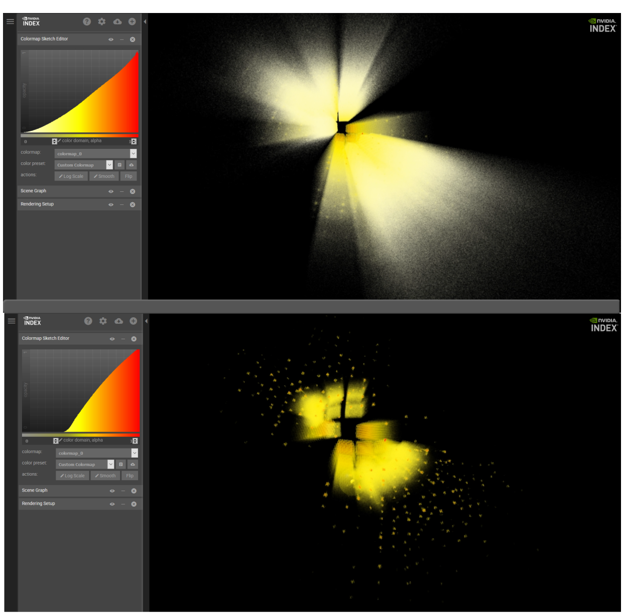

Figure 4: A demonstration of how color of IndeX visualizations can be adjusted to view features of interest.

However, effective data visualization and exploration must be interactive. Long waits between user interactions and their results will be frustrating and waits greater than one second can easily break a train of thought [11]. Consequently, it is necessary to ensure that IndeX is interactive under real-world conditions.

Interactivity Benchmarks We quantify IndeX's interactivity using two metrics, frames per second (FPS) and playback delay. FPS is a measure of how smoothly interactions are shown, while playback delay measures how long users have to wait before the results of their interactions are visible. We measure these values over a one minute period on various quality network connections and display both the average and "worst" values of these metrics. The poorest values are typically achieved by rapid user interaction at a high zoom level. Figure 5 shows that, on a modern mobile network connection, even the worst achievable playback delay is lower than one second, and the average playback delay is never higher than $1/4^{th}$ of a second, even on the worst connections.

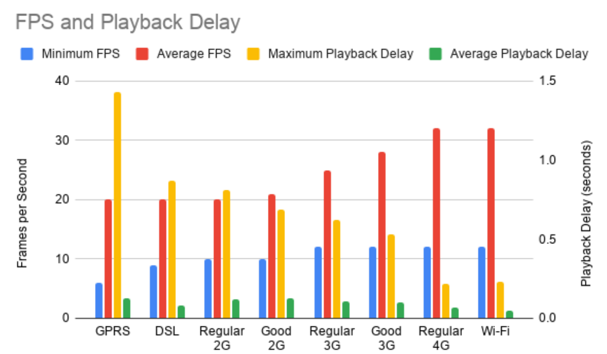

Figure 5: The FPS and playback delay of IndeX on varying networks connections. The average and "worst" values are shown, computed by recording a minute of user interaction

Jupyter Notebook Integration IndeX can also be displayed in a Jupyter Notebook, which provides the ability for users to create a custom UI adjacent to the IndeX UI, as well as display the IndeX visualization alongside their own analyses. An example showing IndeX in a Jupyter Notebook with a custom UI alongside a Bragg peak analysis can be seen in Fig. 2.

CONCLUSION

We have demonstrated that IndeX is an effective tool for visualizing TOPAZ datasets and fills a gap in the capabilities of current neutron scattering analysis packages. It provides high quality, interactive visualizations even over poor network connections. Additionally, we have shown that IndeX integrates well with Jupyter Notebook, a common data analysis platform.

We have also demonstrated that Open Voxel Data Bases are an effective and space efficient way to store TOPAZ data for visualization, reducing the size of the data by up to five orders off magnitude, allowing us to visualize more, larger datasets.

ISBN: 978-3-95450-214-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2021-TUPAB413

REFERENCES

[1] A. J. Schultz *et al.*, "Integration of neutron time-of-flight single-crystal Bragg peaks in reciprocal space", *Journal of Applied Crystallography*, vol. 47, pp. 915-921, Jun. 2014. doi:10.1107/S1600576714006372

[2] G. Jogl *et al.*, "High resolution neutron crystallographic studies of the hydration of coenzyme cob(11)alamin", *Acta Crystallographica Section D*, vol. 67, pp. 584-591, Jun. 2011. doi:10.1107/S090744491101496X

[3] L. Coates *et al.*, "A suite-level review of the neutron single-crystal diffraction instruments at Oak Ridge National Laboratory", *Review of Scientific Instruments*, vol. 89, p. 092802, 2018. doi:10.1063/1.5030896

[4] J. Weng *et al.*, "K-space algorithmic reconstruction (KAREN): a robust statistical methodology to separate Bragg and diffuse scattering", *Journal of Applied Crystallography*, vol. 53, pp. 159-169, Jan. 2020. doi:10.1107/S1600576719017060

[5] O. Arnold *et al.*, "Mantid-Data Analysis and Visualization Package for Neutron Scattering and μ SR Experiments", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 764, pp. 156–166, Nov. 2014. doi:10.1016/j.nima.2014.07.029

[6] T. M. Michels-Clark *et al.*, "Expanding Lorentz and Spectrum Corrections to Large Volumes of Reciprocal Space for Single-Crystal Time-of-Flight Neutron Diffraction". *Journal of Applied Crystallography*, vol. 49, pp. 497-506, 2016. doi:10.1107/S1600576717006781

[7] A. Henderson, *The ParaView Guide: A Parallel Visualization Application*, Clifton Park, NY, USA: Kitware Inc., 2007.

[8] NVIDIA IndeX® 3D Volumetric Visualization Framework, https://developer.nvidia.com/nvidia-index

[9] NVIDIA IndeX Plugin in ParaView 5.5, https://blog.kitware.com/nvidia-index-plugin-in-paraview-5

[10] K. Museth, "VDB: High-Resolution Sparse Volumes with Dynamic Topology", *ACM Transactions on Graphics*, vol. 32, no. 3, pp. 1-22, Jun. 2013. doi:10.1145/2487228.2487235

[11] J. A. Hoxmeier and C. DiCesare, "System Response Time and User Satisfaction: An Experimental Study of Browser-based Applications", in *Proc. of the Association of Information Systems Americas Conf.*, Long Beach, CA, USA, Aug. 2000, pp. 140-145. https://aisel.aisnet.org/amcis2000/347