Self-driving Multimodal Studies at User Facilities

Anonymous Author(s)
Affiliation
Address
email

Abstract

Multimodal characterization is commonly required for understanding materials. User facilities possess the infrastructure to perform these measurements, albeit in serial over days to months. In this paper, we describe a unified multimodal measurement of a single sample library at distant instruments, driven by a concert of distributed agents that use analysis from each modality to inform the direction of the other in real time. Powered by the *Bluesky* project at the National Synchrotron Light Source II, this experiment is a world's first for beamline science, and provides a blueprint for future approaches to multimodal and multifidelity experiments at user facilities.

1 Introduction

Fully characterizing new materials depends on multiple modalities of measurement, e.g., spectroscopy and diffraction. Differing modalities often suffer from contrasting fidelity and resource requirements. A common example of this is the use of X-ray absorption fine structure (XAFS) and total scattering for the characterization of functional materials such as high entropy alloy [1], high entropy oxides [2], and electroactive materials [3]. Probing these modalities simultaneously and efficiently would accelerate materials analysis, and by extension discovery.

In many cases, these analysis techniques are only available to researchers at light sources and user facilities. Light sources, such as the National Synchrotron Light Source II (NSLS-II) at Brookhaen National Laboratory (BNL), are large scale facilities that provide service science for the research community. These are government funded centers that possess first or only in class measurement capabilities. A *beamline* or *end-station* is the instrumentation that provides a measurement capability at a light source. Due to advances in both light source accelerator and detector technologies, the productivity of a high-throughput beamline is no longer limited by the amount of photons it can produce and detect, but rather the ability to control and analyze the high rate of measurements. To help realize and leverage the full potential of these facilities, researchers have automated data collection and integrated artificial intelligence (AI) into the real-time analysis and orchestration of experiments [4, 5, 6]. In line with this, recent advancements have been made to convert and incorporate beamlines into self-driving labs or materials acceleration platforms (MAPs) [7, 8, 9, 10].

Multimodal measurements are a critical part of NSLS-II capabilities, yet present a further challenge: since time is allocated using a proposal system, a sample whose measurement warrants further study using a different modality may have to wait months after analysis to be allocated time on the next instrument. This bottleneck is exemplified by the *in operando* study of a lithium–sulfur battery cell at three different beamlines [3]. Moreover, the diversity of beamlines and materials studied creates a need for diverse agents to interface with experimental orchestration [10, 8]. Integrating multiple beamlines using distributed agents would not only open the bottleneck for the analysis of engineered devices or combinatorial materials [11], but enable discovery workflows with self-driving synthesis engines.

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

In this work, we demonstrate the world's first truly multimodal measurement at a light source. The measurement used two physically distant beamlines simultaneously to examine a single sample library, in a decentralized control loop. Multiple AI agents were able to guide both beamlines in concert, while retaining the opportunity for human experts to engage in decision making. This work can scale readily to incorporate other beamlines and MAPs, as well as serving as a blueprint for building the experimental orchestration of other MAPs.

2 Preliminaries

X-ray powder diffraction X-ray powder diffraction (XRD) is one of the primary characterization methods of the material science and solid state chemistry communities. Diffractometers measure scattered intensity from a sample as a function of angle, from which the atomic-scale details can be extracted such as material phase, lattice constant, strain, site occupancy, and atomic displacement parameters. At large scale light source user facilities, measurement times for full XRD patterns are on the order of millisecond to minutes, dependent on the sample and instrumentation. At the PDF beamline at NSLS-II, it is common to measure a full XRD pattern in 30 seconds [12].

X-ray-Absorption Fine-Structure spectroscopy X-ray-Absorption Fine-Structure spectroscopy (XAFS) is another characterization method in wide use in materials science, solid state chemistry, and many other scientific disciplines. In XAFS, an X-ray beam is scanned in energy over a range that includes the binding energy of deep-core electrons in an element contained in the sample. By measuring the change in X-ray absorption cross section as a function of energy, information about the electronic state of the element as well as the local atomic configuration can be obtained. At the BMM beamline at NSLS-II, an XAFS spectrum is typically measured in about 7 minutes [13].

Measurement task The measurement task herein is to fully and efficiently discover and characterize all material phases in a sample library, where the samples vary as a function of ordinate on a single wafer. As beamtime at user facilities is a limited resource, optimal collection strategies allow for higher duty cycles, ultimately resulting in more samples being measured and thus, more scientific productivity. Therefore we consider the optimal exploration of the space $\mathcal{X}$, where we can probe space *via* two mappings: $f_i : \mathcal{X} \mapsto \mathbb{R}^{n_i}$, where $\mathcal{X}$ is the continuous space of positions in a wafer library Σ, and n is the data length of a given spectral measurement. In each phase of measurement we select k new points $X_{\text{query}} \subset \mathcal{X}$ at which to query f_i. The goal is to find all unique values of f_i with the fewest number of queries.

3 Decentralized control of beamlines

We depended on two broad advancements to enable multimodal analysis MAPs. Firstly, we deployed industry standard technology to limit isolation, improve system reliability, and scale effectively. Secondly, we focused on moving beyond closed loop control for experiment orchestration. These advancements apply to almost all 28 operational beamlines at NSLS-II, with 22 more expected to be built.

Infrastructure developments for scalability A recent multi-million dollar effort, BNL has built a *High Throughput Science Network* that connects NSLS-II to BNL super computers and research centers [10]. This included a reprovisioning of the IP space of NSLS-II that enabled firewalled communication between beamlines, the data storage center, and compute for data processing. Virtualization was used to deploy all services in this work using VMWare clusters run centrally and at the beamlines [14]. This enabled increased fault tolerance and rapid prototyping of tools. Redundant MonogDB [15] services were deployed for document style data storage and retrieval, with a central Lustre [16] file store for most large data. Data access was largely achieved using Tiled [17], which is a web-based storage agnostic tool replacing DataBroker [18]. A Kafka service [19] was deployed that published the document stream of all beamline measurements, which were then subscribed to by agents or other services. At a facility scale the Ansible automation platform is used to scale provisioning of VMs and services [20]. Lastly, secure remote access to beamlines was acheived using Guacamole [21], with data access and remote analysis provided by a JupyterHub [22].

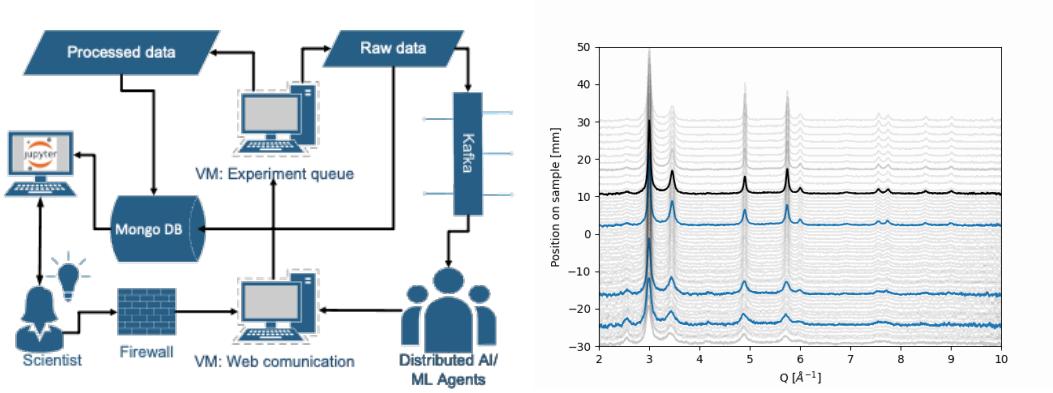

Figure 1: Left: block diagram for decentralized orchestration of a single beamline. Each beamline can be linked *via* agents with access to the respective VM for web communication (http-server), Kafka node, and/or database. Right: a snapshot of the commissioning experiment. The PDF monarch read all historical XRD data (grey), suggested the next point in a geometric design of experiments for XRD (black), and periodically suggested regions of maximal explained variance for further study using XAFS (blue).

Experimental orchestration using the *Bluesky* suite The *Bluesky* project is a collection of python libraries for experimental science with thousands of users and nearly 100 community developers working on active forks. While co-developed, each package is designeed to be independently used [18]. A beamline consists of many devices (motors, detectors, pumps, sensors, etc.) that can be orchestrated to conduct experimental plans as asynchronous coroutines using the *Bluesky Run Engine*. Current 'autonomous experiments' at user facilities operate in synchronous closed loop measurements over a single modality [6]. This lock-step approach to experiment and analysis, leaves no room for human experts to engage in the loop, or the incorporation of information from complementary techniques.

With recently developed packages in the *Bluesky* project and the infrastructure advancements at NSLS-II, we first executed decentralized control of a single beamline using AI and human agents. A queue server [23] replaced the command line execution of plans by the *Run Engine* with a mutable priority queue. The queue server manages permissions, and was securely accessed by human and AI agents through http protocols using the http-server [24], and monitored through a graphical user interface. A dedicated queue server was run on a VM within a beamline subnet, with dedicated http-servers run on VMs on a central network. As the beamline executed plans, it published raw data to a Kafka topic and wrote to a central database. Services subscribed to the raw data stream and produced processed data that was concurrently stored and referenced in the database. Agents then accessed the raw or processed data to produce visualizations or suggest follow up experimental plans using the http-server. There was no limit to the number of agents, or their location. In this work, AI agents were deployed on physical machines within the NSLS-II IP space, and human agents engaged remotely.

4 Agent driven multimodal characterization

In this work, we combined the decentralized experimental orchestration of two distant beamlines *in concert*. Together, both beamlines were exploring an identical sample library, seeking to extract the most information from the library. The control system included passive agents for data processing, unsupervised agents that combined a design of experiments approach and entropy search, pretrained deep ensembles, and human expert agents. This simultaneous combination of beamlines demonstrates a world first for synchrotron science, and lays the framework for future experiments using more specialized agents and complex experimental spaces.

Agent design The agents designed for this experiment considered what was rational, time sensitive, and interesting. The agents were developed using an `ask-tell-report` paradigm, where an agent could be 'told' about new data, 'asked' what to do next, and 'report' about it's current status [5].

Each step was recorded using the streaming event model of *Bluesky*, *via Tiled*, so that the agent's perspective and decision making could be played back for investigation.

We used a monarch–subject relationship to let an agent subscribing to one beamline dictate the plans of the opposite beamline. At baseline, an agent would propose a design of experiments approach for a measurement on the queue. The monarch–subject agents gave the "monarch" agent priority to dictate the next measurements on the "subject" queue at some interval. For example, the PDF-monarch consumed all the XRD data measured at PDF, and every hour would dictate some regions of interest for further analysis at the BMM-subject. Because the PDF measurement was much faster than BMM, the BMM-monarch would dictate at every round of design. Regions of interest were determined *via* maximal explained variance using alternating K-means clustering and constrained matrix factorization [25]. Passive agents were processing and visualizing all data from the experiments for expert interpretation. Human experts were also able to perform real-time analysis using the JupyterHub to gain their own insights, and add suggested plans to the queue.

Additional AI agents were built for the experiment but were unused. These included deep feed-forward ensemble models that consumed 1-d spectra and predicted the corresponding crystallographic phase, and convolutional variational autoencoders to compress the spectral space into a lower dimensional subspace [26, 27]. These models were pretrained on synthetic data produced using the computational equilibrium phase diagram from the Materials Project [28]. Still using the monarch–subject paradigm, the ensemble model was designed to suggest regions of maximum information entropy. To supersede design of experiments, a Gaussian process was continually retrained to predict the latent representation of the VAE, and be sampled with Bayesian optimization using an entropy search. In practice, the sample library that was synthesized did not reflect the full phase diagram used to train these deep models, thus only the naive and unsupervised models were used in the commissioning study.

Commissioning example For this initial commissioning study, we selected a sample libray that was both readily characterized by the two techniques and assuredly identical in composition. To accomplish the former, we cast a bimetallic alloy (CuTi) onto a circular Si wafer, where the variable composition was dependent on the position on the wafer (i.e. high concentration of Cu on one side, high concentration of Ti on the other, linearly varying across the wafer). Although the fabrication is reproducible, we ensured identical samples by cleaving the prepared circular wafer in half, parallel to the direction of the deposition gradient. Thus, the composition of the sample could be considered functionally identical at either half of the cleave. The wafers were mounted on standard stages on both beamlines, such that they could be independently positioned with the instrument stage to any point within 1-5 seconds.

5 Discussion and Future Work

While our initial commissioning experiment demonstrates a unique capability for beamline science, there are several avenues for improvement. From an infrastructure standpoint, services that were deployed using virtualization (e.g., the queue-server), could be reproducibly developed, version controlled, and efficiently deployed using containerization; however, this would require additional expertise. Furthermore, integrating synthesis MAPs that are not natively developed using *Bluesky*, would unlock on-the-fly synthesis and characterization of dynamic sample libraries. To this end, more complex experimental spaces will be explored in the future, which will in turn warrant the integration of more complex agents. Specifically, methodological work is needed to build physics-aware agents that can effectively combine both the XRD and XAFS data, and use information from both measurements to drive both measurements.

In this work we demonstrated the capability and blueprint to conduct self-driving multimodal analysis using multiple beamlines. Using decentralized control of two self-driving beamlines, we allowed for multiple agents—both computational and human—to contribute to the experimental planning. We commissioned this technology in the study of a well understood bimetallic alloy library, with a more complex library scheduled for study in November 2022. This work opens up new opportunities for high-throughput user science that studies more complex multimodal and multifidelity tasks, and the development of physics-imbued multimodal AI agents. We expect facile incorporation of new beamlines at NSLS-II or other facilities using *Bluesky* (e.g., the combined neutron and X-ray measurement of a system). This world's first marks a step forward in making optimal use of light sources and user facilities.

References

[1] D. B. Miracle and O. N. Senkov. A critical review of high entropy alloys and related concepts. *Acta Materialia*, 122:448–511, 2017. doi: https://doi.org/10.1016/j.actamat.2016.08.081. URL http://www.sciencedirect.com/science/article/pii/S1359645416306759.

[2] Brianna L. Musicó, Dustin Gilbert, Thomas Zac Ward, Katharine Page, Easo George, Jiaqiang Yan, David Mandrus, and Veerle Keppens. The emergent field of high entropy oxides: Design, prospects, challenges, and opportunities for tailoring material properties. *APL Materials*, 8(4): 040912, 2020. doi: 10.1063/5.0003149. URL https://doi.org/10.1063/5.0003149.

[3] Ke Sun, Chonghang Zhao, Cheng-Hung Lin, Eli Stavitski, Garth J. Williams, Jianming Bai, Eric Dooryhee, Klaus Attenkofer, Juergen Thieme, Yu-chen Karen Chen-Wiegart, and Hong Gan. Operando multi-modal synchrotron investigation for structural and chemical evolution of cupric sulfide (cus) additive in li-s battery. *Scientific Reports*, 7(1):12976, 2017. doi: 10.1038/s41598-017-12738-0. URL https://doi.org/10.1038/s41598-017-12738-0.

[4] Stuart I Campbell, Daniel B Allan, Andi M Barbour, Daniel Olds, Maksim S Rakitin, Reid Smith, and Stuart B Wilkins. Outlook for artificial intelligence and machine learning at the NSLS-II. *Machine Learning: Science and Technology*, 2(1):013001, mar 2021. doi: 10.1088/2632-2153/abbd4e. URL https://doi.org/10.1088/2632-2153/abbd4e.

[5] Tatiana Konstantinova, Phillip M. Maffettone, Bruce Ravel, Stuart I. Campbell, Andi M. Barbour, and Daniel Olds. Machine learning enabling high-throughput and remote operations at large-scale user facilities. *Digital Discovery*, pages –, 2022. doi: 10.1039/D2DD00014H. URL http://dx.doi.org/10.1039/D2DD00014H.

[6] Marcus M. Noack, Petrus H. Zwart, Daniela M. Ushizima, Masafumi Fukuto, Kevin G. Yager, Katherine C. Elbert, Christopher B. Murray, Aaron Stein, Gregory S. Doerk, Esther H. R. Tsai, Ruipeng Li, Guillaume Freychet, Mikhail Zhernenkov, Hoi-Ying N. Holman, Steven Lee, Liang Chen, Eli Rotenberg, Tobias Weber, Yannick Le Goc, Martin Boehm, Paul Steffens, Paolo Mutti, and James A. Sethian. Gaussian processes for autonomous data acquisition at large-scale synchrotron and neutron facilities. *Nature Reviews Physics*, 3(10):685–697, 2021. doi: 10.1038/s42254-021-00345-y. URL https://doi.org/10.1038/s42254-021-00345-y.

[7] Phillip M Maffettone, Joshua K Lynch, Thomas A Caswell, Clara E Cook, Stuart I Campbell, and Daniel Olds. Gaming the beamlines—employing reinforcement learning to maximize scientific outcomes at large-scale user facilities. *Machine Learning: Science and Technology*, 2 (2):025025, mar 2021. doi: 10.1088/2632-2153/abc9fc. URL https://doi.org/10.1088/2632-2153/abc9fc.

[8] Andi Barbour, Stuart Campbell, Thomas Caswell, Masafumi Fukuto, Marcus Hanwell, Andrew Kiss, Tatiana Konstantinova, Ricarda Laasch, Phillip Maffettone, Bruce Ravel, and Daniel Olds. Advancing discovery with artificial intelligence and machine learning at nsls-ii. *Synchrotron Radiation News*, 0(0):1–7, 2022.

[9] Martin Seifrid, Jason Hattrick-Simpers, Alán Aspuru-Guzik, Tom Kalil, and Steve Cranford. Reaching critical mass: Crowdsourcing designs for the next generation of materials acceleration platforms. *Matter*, 5(7):1972–1976, 2022/07/12 2022. doi: 10.1016/j.matt.2022.05.035. URL https://doi.org/10.1016/j.matt.2022.05.035.

[10] Brookhaven National Laboratory. 2022 NSLS-II Strategic Plan. Technical report, National Synchrotron Light Source II, https://www.bnl.gov/nsls2/docs/pdf/nsls2-strategic-plan.pdf, September 2021.

[11] Martin L. Green, Ichiro Takeuchi, and Jason R. Hattrick-Simpers. Applications of high throughput (combinatorial) methodologies to electronic, magnetic, optical, and energy-related materials. *Journal of Applied Physics*, 113(23):231101, 2013. doi: 10.1063/1.4803530. URL https://doi.org/10.1063/1.4803530.

[12] Brookhaven Science Associates. *Pair Distribution Function*, 2022. https://www.bnl.gov/nsls2/beamlines/beamline.php?r=28-ID-1 [Accessed: 2022-09-26].

[13] Brookhaven Science Associates. *Beamline for Materials Measurement*, 2022. https://www.bnl.gov/nsls2/beamlines/beamline.php?r=6-BM [Accessed: 2022-09-26].

[14] D. Kusnetzky. *Virtualization: A Manager's Guide*. O'Reilly, 2011.

[15] Kristina Chodorow and Michael Dirolf. *MongoDB: The Definitive Guide*. O'Reilly Media, Inc., 1st edition, 2010. ISBN 1449381561.

[16] Steve Pate. *UNIX Filesystems*. Veritas. John Wiley & Sons, Nashville, TN, January 2003.

[17] Brookhaven National Laboratory. Tiled. https://github.com/bluesky/tiled, 2022.

[18] Daniel Allan, Thomas Caswell, Stuart Campbell, and Maksim Rakitin. Bluesky's ahead: A multi-facility collaboration for an a la carte software project for data acquisition and management. *Synchrotron Radiation News*, 32(3):19–22, 2019. doi: 10.1080/08940886.2019.1608121. URL https://doi.org/10.1080/08940886.2019.1608121.

[19] Neha Narkhede, Gwen Shapira, and Todd Palino. *Kafka: The Definitive Guide Real-Time Data and Stream Processing at Scale*. O'Reilly Media, Inc., 1st edition, 2017. ISBN 1491936169.

[20] Ansible Project. *Anisible Documentation*, 2022. https://docs.ansible.com/ansible/latest/index.html [Accessed: 2022-09-26].

[21] Tara Seals. Apache guacamole opens door for total control of remote footprint, 2020. https://threatpost.com/apache-guacamole-control-remote-footprint/157124/ [Accessed: 2022-09-26].

[22] Michael Milligan. Interactive hpc gateways with jupyter and jupyterhub. In *Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact*, PEARC17, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450352727. doi: 10.1145/3093338.3104159. URL https://doi.org/10.1145/3093338.3104159.

[23] Brookhaven National Laboratory. bluesky-queueserver. https://github.com/bluesky/bluesky-queueserver, 2022.

[24] Brookhaven National Laboratory. bluesky-httpserver. https://github.com/bluesky/bluesky-httpserver, 2022.

[25] Phillip M. Maffettone, Aidan C. Daly, and Daniel Olds. Constrained non-negative matrix factorization enabling real-time insights of in situ and high-throughput experiments. *Applied Physics Reviews*, 8(4):041410, 2021. doi: 10.1063/5.0052859. URL https://doi.org/10.1063/5.0052859.

[26] Phillip M. Maffettone, Lars Banko, Peng Cui, Yury Lysogorskiy, Marc A. Little, Daniel Olds, Alfred Ludwig, and Andrew I. Cooper. Crystallography companion agent for high-throughput materials discovery. *Nature Computational Science*, 1(4):290–297, 2021.

[27] Lars Banko, Phillip M. Maffettone, Dennis Naujoks, Daniel Olds, and Alfred Ludwig. Deep learning for visualization and novelty detection in large x-ray diffraction datasets. *npj Computational Materials*, 7(1):104, 2021.

[28] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A. Persson. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, 2013.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes] We kept the claims centered on the infrastructure and experimental capacity, without substantial model based claims.
   (b) Did you describe the limitations of your work? [Yes] First paragraph of discussion.
   (c) Did you discuss any potential negative societal impacts of your work? [N/A]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] Appendix Code availability
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] XCA scripts are available in attached repository
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [No] See referenced XCA publications (Maffettone 2021, Banko 2022)
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes]
   (b) Did you mention the license of the assets? [Yes] Appendix Code availability
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [Yes] BSD 3
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

A Code availability

All of the code for driving the experiments and developing the models is available at github.com/NSLS-II-PDF/mmm-experiments under the BSD 3-Clause license. This is a repository under active development, with the last commit in line with the described work at 2299acc11bb911e72e68e4b658d022cc775b5868.

Similarly, the underlying components are developed under the BSD 3-Clause license as parts of the *Bluesky* project. For security reasons, database and network configurations are not publically available.

B Agent training

Details on training the deep ensemble models and VAE are available in the above repository. All models were trained continuously using live data synthesis for 1 week prior to the experiment on an NVIDIA A1000 GPU. The agents were deployed on a single workstation with a 32 core CPU during the experiment.

C Naive agent selections

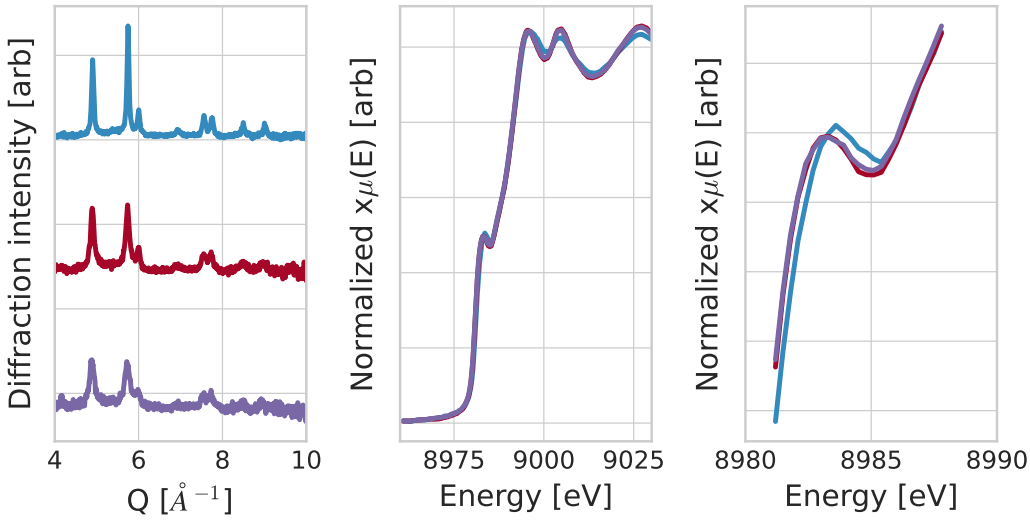

Figure 2: Stacked plots of the positions suggested by the agent in Figure 1. The changes between the regions with increasing Cu content were extremely subtle, which drove the consideration of unsupervised techniques over pretrained deep models that were conditioned on more phases. Left: The XRD of the three phases over a Q-range that shows some short range order changes in the patterns. Middle: The XAFS spectra showing the low energy fine structure. Right: The leading edge of the spectra, also known as the X-ray absorption near edge structure (XANES).