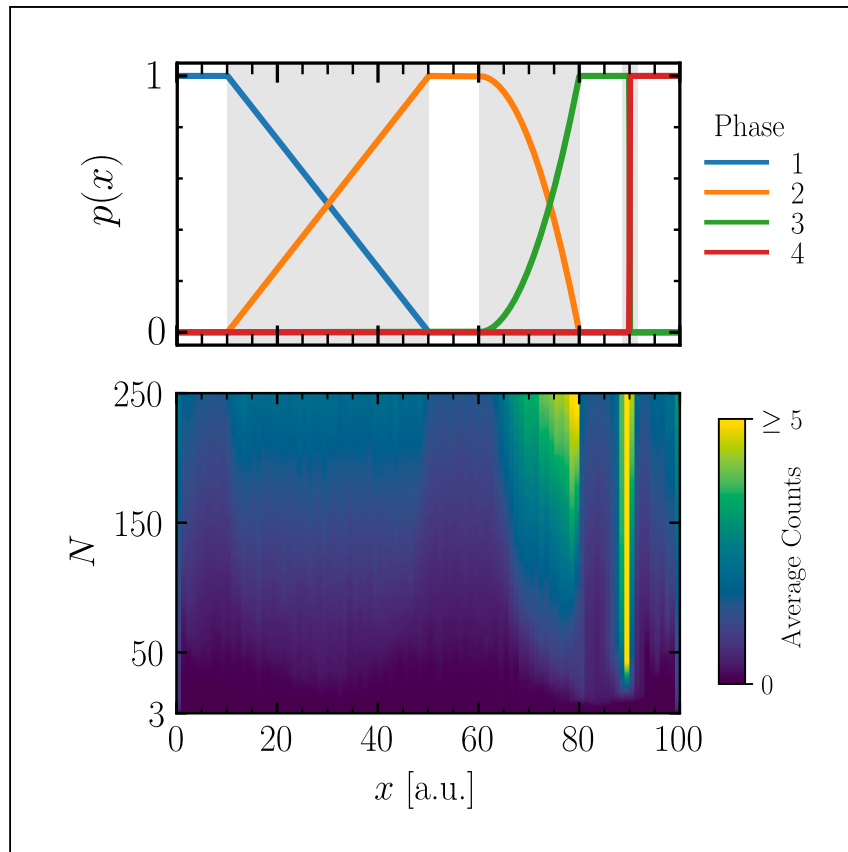


Article

Flexible formulation of value for experiment interpretation and design



Driving the characterization of materials to focus on regions where phases are changing the fastest.

Matthew R. Carbone, Hyeong Jin Kim, Chandima Fernando, ..., Bruce Ravel, Yugang Zhang, Phillip M. Maffettone

mcarbone@bnl.gov (M.R.C.)
yuzhang@bnl.gov (Y.Z.)
pmaffetto@bnl.gov (P.M.M.)

Highlights

A new technique for objective-free optimization and epistemic research problems

Extensibility to diverse experiments and model inclusion

Compatible with optimization tools, such as Bayesian optimization



Understanding

Dependency and conditional studies on material behavior

Carbone et al., Matter 7, 685–696
February 7, 2024 © 2023 Elsevier Inc.
<https://doi.org/10.1016/j.matt.2023.11.012>



Article

Flexible formulation of value for experiment interpretation and design

Matthew R. Carbone,^{1,*} Hyeon Jin Kim,² Chandima Fernando,³ Shinjae Yoo,¹ Daniel Olds,³ Howie Joress,⁴ Brian DeCost,⁴ Bruce Ravel,⁴ Yugang Zhang,^{2,*} and Phillip M. Maffettone^{3,5,*}

SUMMARY

The challenge of optimal design of experiments pervades materials science, physics, chemistry, and biology. Bayesian optimization has been used to address this challenge but requires framing experimental campaigns through the lens of maximizing some observable. However, this framing is insufficient for epistemic research goals that seek to comprehensively analyze a sample space, without an explicit scalar objective. In this work, we propose a flexible formulation of scientific value that recasts a dataset of input conditions and higher-dimensional observable data into a continuous, scalar metric. Intuitively, the function measures where observables change significantly, emulating the perspective of experts driving an experiment. We demonstrate this as a collaborative analysis tool and objective for optimization technique using two simulated and two experimental examples. The method is flexible, easily deployed, seamlessly compatible with existing optimization tools, can be extended to multi-modal and multi-fidelity experiments, and can integrate existing models of an experimental system.

INTRODUCTION

The combination of automation and artificial intelligence (AI) to create closed-loop self-driving, autonomous laboratories—or human-interfaced acceleration platforms—has begun revolutionizing scientific research across chemistry,^{1,2} biology,^{3,4} and materials science.^{5–9} These contemporary platforms mostly use single AI agents, but can also leverage the added value of multiple agents working in tandem.¹⁰ To date, most efforts in agent development have focused on designing algorithms that optimize a target value¹¹ or resource allocation.^{12,13} Unfortunately, these approaches to agent design do not encompass the research settings where the objective is more epistemic; that is, the research pertains to comprehensive understanding or interpretation of an experimental space,¹⁴ and not the optimization of a target. Examples of epistemic objectives are ubiquitous in characterization,¹⁵ user facilities, and “science as a service” platforms.¹⁶ These bring forth a new challenge in how to leverage AI advancements for optimal experiment design.

Research motivated by comprehensive understanding of a system is common across disciplines. It appears in problems involving spatial characterization or fixed sample spaces, such as device mapping,^{17,18} tomography,⁷ or phase mapping.^{19,20} It also is recurrent when searching large plateaus of space for sharp changes, such as in searching for reactivity²¹ or phase changes.^{20,22} Nonetheless, when research questions are directed more by understanding a system than by optimizing the system for a single property, certain measurements will still prove more valuable than others. Traditional

PROGRESS AND POTENTIAL

Scientific experiments, particularly those at large-scale user facilities, are increasingly automated. With the advent of new technologies for optimal experimental design, automated experiments are being pushed to higher degrees of autonomy. However, many approaches to self-driving experiments are focused on optimizing some objective or seeking novelty. Our research addresses the challenge of experimental design for epistemic goals. We developed a formulation of value for arbitrary types of measurements, which enables researchers from broad domains to tackle the sole objective of understanding their systems. This is a pressing need for facilities that provide characterization services for diverse samples. The method is compatible with optimization tools and provides a potent analysis technique. Given the common challenge of formulating a scalar metric or objective to optimize, we believe our research could be used as a default agent in materials acceleration platforms.



design of experiment (DOE) approaches the exploration of a known space to explain the variation of a response function in that space, albeit it is not adaptive and expects each input parameter to impact the response function.²³ In the case of allocating limited resources over independent samples, reinforcement learning (RL) has been used to extract maximal value.¹² When a model for the system is available, Bayesian inference can be used to query data that will best mitigate the uncertainty of that model.^{24–26} Bayesian optimization over expected information gain²⁷ can be used to explore the experimental space; however, without a conversion between an observable and a finite objective, Bayesian optimization cannot be effectively leveraged.

There is also a significant body of work examining alternative approaches to the design problem, particularly in the context of RL and rewarding discovery.²⁸ “Curiosity-driven” robots,^{29,30} and more generally knowledge-based methods,^{31,32} have been used for self-driving and exploratory experiments, with expectations of foundation models aiding scientific exploration.³³ Other works suggest abandoning objectives and rewards entirely, positing that novelty alone is sufficient to supersede objective-driven evolutionary algorithms³⁴ or for learning new skills in simulated robotic tasks.³⁵ While novelty is a meaningful goal in its own right, many experimental designs—especially at large scientific user facilities—are as concerned with understanding connections between unique observations, as much as the discovery of those novel observations. This is in line with the motivations of our colleagues consulted in this work, and outlined by the sociology of science: specifically, that scientific “awards are linked with strategies more likely to bridge disconnected network components.”³⁶

With an epistemic goal, an optimal agent will therefore yield an experimental design that produces the best dataset for understanding the experimental space. This understanding would be derived from expert interpretation, modeling, AI, or some combination of techniques. Furthermore, the agent should be able to operate with or without a model of space or the observable. It must also be robust to the “cold start” problem,³⁷ operating efficiently under initially extremely data- and information-limited conditions. Last, it would be beneficial for any agent to make use of contemporary advancements in optimization methods.

Herein, we propose a generic scientific value function (SVF) that recasts a dataset of observables into scalar measures of “value” by mirroring the perspective and actions of human experts. While natively model free, the SVF can incorporate models of the experimental or observable space. Crucially, it can be used as an optimization target in other procedures, such as Bayesian or Monte Carlo optimization.²⁷ We demonstrate the application of the SVF through (1) a simulated X-ray diffraction (XRD) phase mapping of first- and second-order transitions; (2) a simulated absorption spectroscopy study of a periodic phase boundary; (3) a variable temperature X-ray total scattering study of BaTiO₃; and (4) an ultraviolet-visible (UV-vis) absorption spectroscopic analysis of nanoparticle synthesis conditions. The adaptive applications of the SVF are accomplished by choosing the next measurement in experimental space according to a Gaussian process³⁸ surrogate model and Bayesian optimization.^{27,39,40} This work creates opportunities for optimal dataset creation and research acceleration without a pre-existing optimization target, and will find broad applicability across scientific disciplines.

RESULTS AND DISCUSSION

A surrogate function for scientific value

We set out to construct a surrogate function for scientific value that would emulate the judgment of an expert scientist without a model for their experimental system.

¹Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11973, USA

²Center for Functional Nanomaterials, Brookhaven National Laboratory, Upton, NY 11973, USA

³National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY 11973, USA

⁴Materials Measurement Science Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

⁵Lead contact

*Correspondence: mcarbhone@bnl.gov (M.R.C.), yuzhang@bnl.gov (Y.Z.), pmaffetto@bnl.gov (P.M.M.)

<https://doi.org/10.1016/j.matt.2023.11.012>

Consider a prototypical example of mapping the phase diagram of a material over multiple dimensions. A rational scientific goal would be to measure every unique phase at least once, and measure with greater resolution across phase boundaries. At the start of a campaign, the measurement of every location has the same potential value. As the campaign progresses, measuring the regions where the observable is not changing with the ordinate becomes less valuable than measuring regions of rapid change. We consulted with staff and users at NSLS-II to understand this intuition, and defined the SVF, U , to capture these behaviors while remaining flexible to include experiment-specific knowledge.

First, we consider an input space $\mathcal{X}$, where queries of $\vec{x}_i \in \mathcal{X}$ comprise a dataset $\mathcal{D}_N = \{(\vec{x}_1, \vec{y}_1), \dots, (\vec{x}_N, \vec{y}_N)\}$, where $\vec{y}_i$ are noisy, multi-dimensional observations of some function, $f(\vec{x}_i)$, such as a diffraction image. We further define two correlation functions for both the input space and the observation space, $h(\vec{x}_i, \vec{x}_j)$ and $g(\vec{y}_i, \vec{y}_j)$, respectively. The default correlation function used in this work for both h and g is the Euclidean distance, or L_2 norm, $\|\cdot\|_{L_2}$. Thus, we define the dataset-dependent SVF as follows:

$$U(\mathbf{x}_i, \mathbf{y}_i, \mathcal{D}_N) = \sum_{(\vec{x}_j, \vec{y}_j) \in \mathcal{D}_N} g(\vec{y}_i, \vec{y}_j) \exp \left\{ -\frac{1}{2} \frac{h(\vec{x}_i, \vec{x}_j)^2}{h_{\min}(\vec{x}_i, \mathcal{D}_N)} \right\}, \quad (\text{Equation 1})$$

where $h_{\min}(\mathbf{x}_i, \mathcal{D}_N)$ is the distance between $\mathbf{x}_i$ and its nearest neighbor in $\mathcal{D}_N$. Using Equation 1, the scientific value can be computed for all inputs in $\mathcal{D}_N$.

The SVF considers the individual value of a new datum with respect to each existing datum and sums over all members of the dataset for a net value. The first term considers where the observable is distinct from those contained in the dataset, and thus valuable. The second term decays that value with respect to how far the data are in the input space, and takes on values between 0 and $e^{-1/2}$ for any $\mathbf{x}_i \in \mathcal{D}_N$. In order to avoid overestimating the value of local regions, the second term is regularized by the nearest neighbors of points in input space. The dynamic feature of this regularization is highlighted in Figures S8 and S11. We considered other forms of the SVF that would use these correlation functions (e.g., proportionate or derivative-like functions, or ones similar to that of Equation 1 but with constant values for $h_{\min}$ [Figure S15]). However, we chose the form of the second term such that it would have a bounded range, and evolving regularization.

This formalism offers a few key features and advantages. First, it adequately reflects the intuition of researchers in practice. It also reduces the dimensionality of the observable space to a scalar objective function that can be readily optimized. While the approach is natively model free, the correlation functions g and h are flexible and can readily incorporate models of the system. For instance, a discrete input space could easily use the Levenshtein or Manhattan distance. With knowledge of the observable space, the distance in a latent space from a variational autoencoder has been used in early implementations of the SVF. Even without a model of the observable space, more involved functions could be considered, such as those from time-resolved pair-correlation functions²⁴ or topological data analysis.²⁵

In the following we made use of Bayesian optimization over a Gaussian process (GP) surrogate model of SVF. Other black-box optimization approaches could be considered, albeit the dataset-dependency of the SVF will constrain their design. The GP used a Matern kernel with homoskedastic noise to construct a probabilistic model

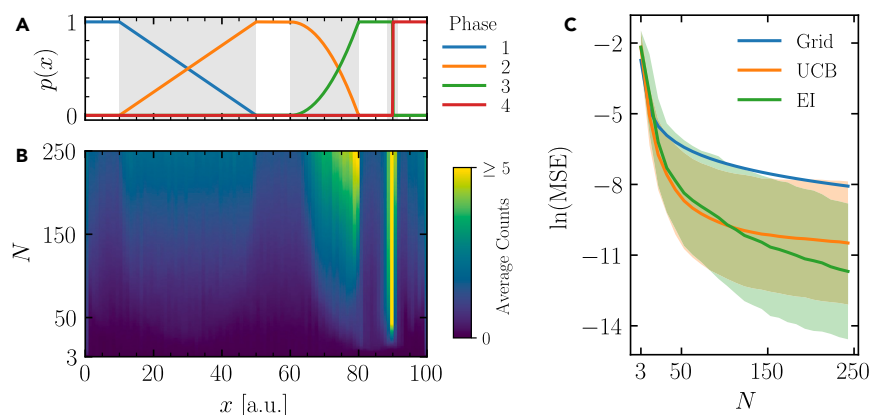


Figure 1. Visual summary of the simulated XRD experiment

(A) The proportion of each of the four phases as a function of position, x . Regions of change are highlighted with a gray background.
(B) A histogram of the average number of counts/experiment as a function of position x and the current dataset size N using the UCB acquisition function with $\beta = 10$. Results are averaged over 300 independent experiments.
(C) The average value of the natural log of MSE as a function of N , plotted with a confidence interval of 2 standard deviations. See also [Figures S1–S8](#).

of U in all input space, including regions where there are no observations. When conditioning the GP, we scaled the value of U to $U \in [0, 1]$. We used the expected improvement (EI)⁴¹ and upper confidence bound (UCB)⁴² acquisition functions for Bayesian optimization. For the UCB functions presented, we used a weighting for variance of $\beta = 10$, although similar results were obtained for values of β ranging from 10 to 100.

We benchmarked this against a common experimental design of measuring over an optimal grid given allotted resources (e.g., time or number of measurements), described here as grid search. We chose this benchmark as the preferred tool of our collaborators, albeit also considered random search strategies ([Figures S2 and S10](#)), and Bayesian inference approaches ([Figures S3 and S12](#)). Herein, we call the SVF modeling procedure used in tandem with the tools of Bayesian optimization (or black-box optimization in general) the Scientific Value Agent (SVA). We note that despite the specific choices used in this work, the extensibility and flexibility of the SVF allows it to be used with any optimization protocol or probabilistic model that can approximate it. While regularization prevents SVA from dwelling on over-sampled regions, we stress the importance of the diversity of thought and models,⁴³ particularly for the potential of combining novelty-seeking approaches⁴⁴ in multi-agent systems.

Characterizing a one-dimensional phase space with simulated XRD

We first tested the SVA *in silico* using the simulated XRD measurement of a library that contained linear mixtures of four phases ([Figure 1](#)). This sampling of a one-dimensional space is common in studying phase behavior over composition or state variables.^{10,17,45} The four XRD patterns corresponding to the phases were defined by a series of randomly placed Gaussian peaks over a constant background ([Figure S1](#)). Normally distributed noise was introduced to the observation at each query by adding a random value to each point of the observation sampled from a distribution with a mean of 0 and standard deviation of 0.1, then squared to ensure positivity. In order to simulate reasonable and challenging types of phase changes, we

chose functional forms to represent first- and second-order transitions: sigmoidal to approximate a discontinuous first-order transition, and linear and quadratic for different rates of second order. These are highlighted in the three shaded regions of [Figure 1A](#). In an optimal measurement of this compositional library, sampling density should be correlated to the rates of change of the phases.

To quantify sampling performance, we tracked the mean squared error (MSE) between the true observation space (phase fractions) and the observation space that could be reconstructed by the sampled queries (assuming oracle knowledge of phase fractions given the observable). The reconstructed dataset is produced by linearly interpolating observations between measured points. As shown in [Figure 1C](#), this metric will decay as more observations are made, with a smaller error corresponding to more robust sampling.

[Figure 1B](#) shows the average sampling histogram of the SVA over 300 independent experiments. Even at small N , we see that the three-phase boundary regions were sampled in proportions commensurate with the rate of change of the phases in those regions. The linear change region was only sparsely sampled (but still sampled more compared with regions of no change), whereas the quadratic region was sampled much more densely. Nonetheless, the region of near-instantaneous change was sampled most densely and earliest, as the algorithm discovered this very sharp boundary, and therefore required more samples to produce an accurate representation of the observable in that region. This sampling density is also reflected in the surrogate model's perceived scientific value over time ([Figure S4](#)).

Both choices of acquisition function outperformed the optimal grid design by roughly an order of magnitude. This performance was apparent in the low- N and limiting cases. It is important to note the balance of exploitation and exploration in these acquisition functions, as attempting to optimize the SVF directly (i.e., greedy optimization) would fail to discover the transition regions ([Figure S5](#)). Not only did the SVA procedures propose experiments in relevant regions of space, they also modeled regions of significant change more efficiently than conventional methods.

We also considered using Bayesian inference as a DOE benchmark. This approach first clustered the data, then trained a probabilistic regressor to predict the cluster labels, and finally queried new points where the uncertainty was maximized.²⁶ We found this methodology to be strongly dependent on the chosen number of clusters, not necessarily more performant than a grid search in the limiting case, and less performant on low N ([Figure S5](#)). While Bayesian inference can be a powerful tool,²⁶ it is intrinsically model dependent, increasing in potency when a more accurate model for the system is available. In Bayesian inference, a label assignment is necessary (here accomplished by K-means clustering). SVA does not require a model or labels, but can incorporate one through the correlation function, $g(\vec{y}_i, \vec{y}_j)$. Considering this comparison and the prevalence of grid searches at actual beamlines, we chose the grid search technique as our benchmark.

Last, we analyzed the technique's robustness with increasingly noisy data ([Figure S6](#)). As the SVF is designed to place value on regions that are highly sensitive to small variations in input, this could lead to an overvaluation of noisy regimes, particularly in the heteroskedastic case where noise changes across the input space. As the signal-to-noise ratio in the simulated data decreases, the correlation term $g(\vec{y}_i, \vec{y}_j)$ would overvalue changes in the noise rather than in the underlying observable. In high-noise situations, a measure of correlation that is not a Euclidean

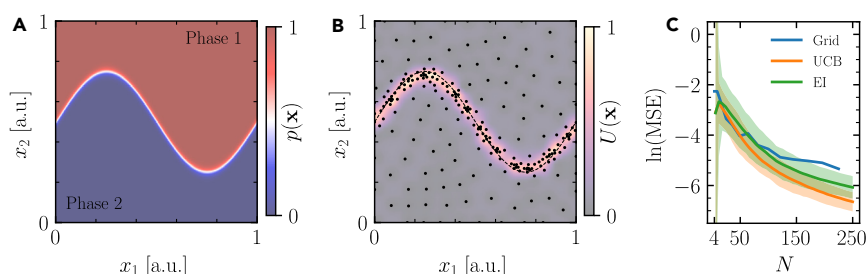


Figure 2. Visual summary of the two-phase, 2-dimensional sinusoidal interface experiment

(A) The two phases, shown in red and blue. The interface, where the phases are in equal contribution, is white.
 (B) The SVF, approximated by a GP and scaled to values between 0 and 1, shown as a backdrop to the sampling results of $N = 250$ points using the UCB acquisition function with $\beta = 10$. The phase boundary is shown as a dotted line.
 (C) The average value of the natural log of MSE as a function of N , plotted with a confidence interval of 2 standard deviations (smaller is better). A total of 300 experiments over random initial points and model/optimizer seeds were performed for experiment statistics. See also [Figures S9–S15](#).

distance between raw observations would be advisable, particularly one that remains invariant under noisy conditions.²²

Characterizing a two-dimensional space with a periodic interface

We completed a second *in silico* test that sought to characterize a two-dimensional library of sample compositions defined by coordinates $\mathbf{x} \in \mathbb{R}^2$. In this case, the library contained only two phases, separated by a sharp periodic boundary ([Figure 2A](#)). The observation of phases and their mixtures was characterized by a spectrum, simulated using Gaussian functions centered at two different locations in space, with noise applied as in the previous example ([Figure S9](#)). As above, phases were linearly mixed, with the proportion of the phases given by a sigmoid function of the position on the wafer,

$$b(x_1) = \frac{1}{2} + \frac{1}{4} \sin(2\pi x_1),$$

$$p(\mathbf{x}) = \frac{1}{1 + \exp\{-50[x_2 - b(x_1)]\}}.$$

(Equation 2)

Designed to be a drastic and challenging test, the resultant phase-dependence on position can be seen in [Figure 2A](#).

We compared the performance of the SVA against conventional methods using the same metrics from above. Again it was clear that a Bayesian optimization approach coupled with SVF outperforms conventional measurement techniques ([Figure 2C](#)). Additionally, we examined how the SVA queried the space around the phase boundary, along with time-resolved plotting of the SVF ([Figure S13](#)) and the density of sampling ([Figure S14](#)). Even in data-limited conditions, the approach successfully mapped out regions of significant change, while still sufficiently sampling relatively constant regions of phase space. As shown in [Figure 2B](#), the sampling focused on the most information-rich region, highlighted around the curve $b(x_1)$.

Compared with the conventional grid search benchmark, the SVA outperformed this baseline by roughly an order of magnitude. The results of a single SVA experiment using the UCB acquisition function with a total of 250 samples show a dense sampling of the interface, without under-sampling the surrounding area. Both UCB and EI behaved

comparably and outperformed the baseline, with UCB being more exploitive of the narrow transition region. Although trade-offs with acquisition functions are expected, optimizing over the SVF was robust regardless of acquisition function choice.

As in the simulated 1-d example, we considered Bayesian inference as a potential baseline for performance (Figure S12). While it is likely that *a priori* knowledge of the two phases or nature of the boundary would improve an exploration using Bayesian inference, using clustering to provide labels to condition a logistic regressor showed poor performance compared with a grid search. The performance was dependent on the chosen number of clusters, and performed worse than the benchmark even when the optimal number of clusters was used.

Characterizing the subtle phase transitions of barium titanate

The final example we present in the active setting highlights the case where naive data-driven approaches fail.⁴⁶ Furthermore, we used this to demonstrate the capacity to integrate more physics-aware correlation functions into the SVF to improve the expressiveness of the surrogate modeling. We emulated a continuous valued experiment in which total scattering data of BaTiO₃ were measured as a function of temperature, by interpolating a dataset measured over 5°C intervals at the pair distribution function beamline at the NSLS-II (Figure S16). These data contain incredibly subtle transitions among four distinct crystallographic phases (rhombohedral, orthorhombic, tetrahedral, and cubic) that are difficult to distinguish using data-driven approaches.⁴⁶ Using established methods, we trained an ensemble of convolutional neural networks to predict these phases from simulated diffraction patterns, and used the trained models to create an encoding of the noisy experimental data.²² We used the SVA procedure to create a surrogate model for an SVF that used these encodings to calculate the observation space correlation function, $g(\vec{y}_i, \vec{y}_j)$.

Following the same procedure as the previous examples, we showcase the results of the sampling as a function of the number of queries in Figure 3B. The SVA correctly identified and attended to the phase transitions extracted by data refinement. We compared approaches by considering the ability of the resultant dataset to construct the Rietveld refined compositions (Figure 3A). Without the inclusion of a deep learned embedding, the SVA produced datasets on par with conventional methods; however, by combining the flexibility of the SVF with a deep learned embedding, it autonomously up-sampled the phase changes of BaTiO₃ (Figure S17).

Because total scattering is a measure of bulk state, it captures more phase coexistence than is present locally throughout the sample, and the first-order phase transitions in BaTiO₃ appear gradual and continuous. As we show in Figure S7, the reconstructive capacity of SVA is on par with the grid search approach when transitions are gradual, but superior when the transitions are sharp. Subsequently, a grid search approach could reconstruct the bulk compositions from Rietveld refinements as well as, if not better than, the SVA approach (Figure 3C). Nonetheless, grid search methods failed to focus on the unique physical behavior of the transitions, which would be exposed through pair distribution function or spectroscopic analysis. This highlights the potential for the SVA to suggest clarifying experiments in the multifidelity or multimodal setting.¹⁰

In-line analysis of nanoparticle synthesis

In the previous three examples, we demonstrated the use of the SVF in an adaptive setting across diverse problems relevant to materials science. We acknowledge that

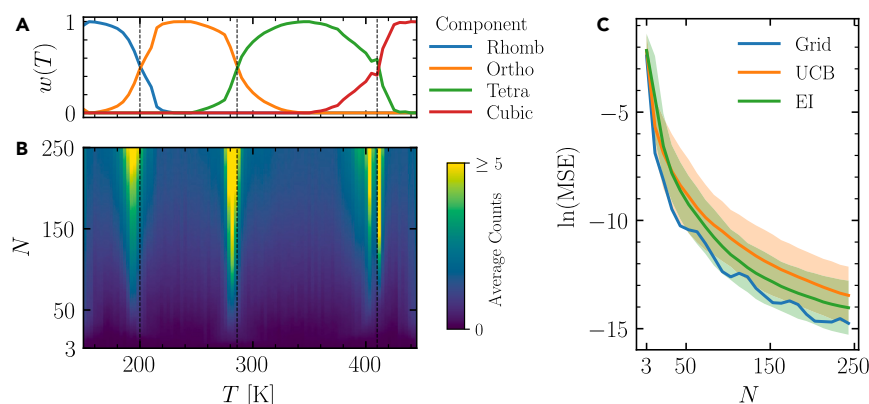


Figure 3. Results of the BaTiO₃ experiment

(A) Phase fractions determined from Rietveld refinement. Refinement results applying Gibbs phase rule to normalized R_{wp} and renormalizing onto $[0, 1]$ (SI). The dotted lines show the detected phase transitions according to the cryostream gas temperature, and will differ by a lag from the sample temperature.

(B) A 2-dimensional histogram of the average number of queries as a function of temperature T and current dataset size using the EI acquisition function. Results are averaged over 300 independent experiments.

(C) The average value of the natural log of MSE as a function of N , plotted with a confidence interval of 2 standard deviations. See also [Figures S16](#) and [S17](#).

a variety of algorithms—or none at all—may be preferable for driving an experiment. Therefore, we highlighted the breadth of the approach by applying the SVF in a passive analysis setting. We deployed the SVF to visualize a spectroscopy dataset produced during an automated nanoparticle synthesis experiment at the Center for Functional Nanomaterials at Brookhaven National Laboratory. The dataset consisted of $N = 375$ experimental flow reactor conditions ($\mathbf{x}_i$) and the corresponding UV-vis absorption spectrum ($\mathbf{y}_i$). The nanoparticle synthesis experiments were performed in a flow reactor by varying four experimental parameters: the volume of sodium citrate (NaCit, 16 mmol/L), chloroauric acid (HAuCl₄, 2 mmol/L), hydrochloric acid (HCl, 10 mmol/L), and sodium hydroxide (NaOH, 10 mmol/L). The total volume of liquid in any experiment is always equal to 40 μL (the size of the droplet in the flow reactor). This reduces the number of degrees of freedom to 3, wherein HCl and NaOH are used to drive the reaction pH. UV-vis absorption spectra were then taken of the final reaction products.

The experiments were performed by domain experts using a grid search with manual intervention. To assist in processing the large dataset of measurements, $\mathcal{D}_N$, we computed $U(\mathbf{x}_i, \mathbf{y}_i; \mathcal{D}_N)$ for all $(\mathbf{x}_i, \mathbf{y}_i) \in \mathcal{D}_N$. This offered the scientists a visual representation of scientific value ([Figure 4A](#)).

The analysis provided by the SVF equipped the scientist with an actionable visualization tool that kept the human in the loop of an otherwise automated experiment. The visualization highlighted regions of high scientific value, as the dataset was growing, by linking the positions in phase space to their individual spectra. By cross referencing the regions of high value with their respective spectra, the scientist was able to examine the SVF analysis and validate it with their own insight. Compared against anomaly detection ([Figure S18](#)), SVF highlights sensitive regions of parameter space instead of outliers in the observable space. For example, the region of [Figure 4B](#) could be interpreted as an anomaly because there is no absorbance, not because the spectrum is changing or interesting. We expect the combination

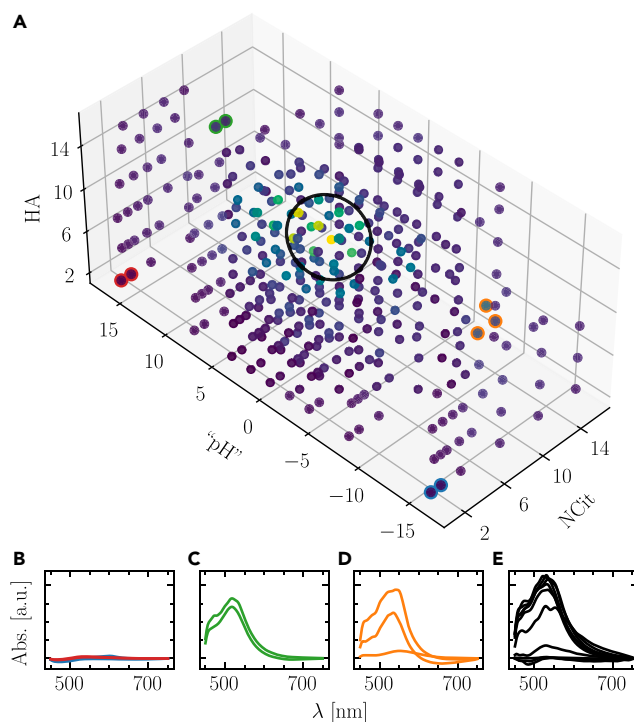


Figure 4. Visual summary of the UV-vis experiment

(A) (Top) The input reactor conditions. Colors are given by the scientific value (scaled to [0, 1], purple to yellow). (Bottom) Spectra corresponding to the colored outlined points in the top plot, shown in order of increasing value as defined by the average SVF in that cluster of points.

(B) The red and blue points show the lowest value (0.01) around static regions of experiment space that produce non-absorbing particles.

(C and D) Increasing patches of scientific value (ranges shown) led the scientist to find (C) green regions (0.09–0.13) with stable and significant absorbance and (D) regions of gradual change in absorbance in orange (0.23–0.35). Last, the SVF analysis shown in black highlights a region of dramatic local change (0.49–0.89). This cluster was selected by taking the point of highest value and finding its 10 nearest neighbors. See also [Figure S18](#).

of this advancement with user interface engineering will create a potent tool that impacts a variety of analysis techniques.

Conclusion

In this work, we presented the SVF, which replicates the judgment of human experts, and recasts a dataset of higher-dimensional observables into scalar measures of value. By quantifying value, the SVF creates an epistemic research objective that can be optimized without the need for feature engineering or a scalar observable. We demonstrated the deployment of the SVF in an adaptive learning context (SVA), how it can be complemented by machine learning or data-reduction techniques, and how it can be used in a streaming analysis deployment for visualizations that accelerate decision making by a human expert. This showed where the SVF can be an important part of a scientist's toolkit, and we expect future opportunities to consider combining SVA with novelty-seeking approaches^{28–32,34} or function optimization^{11–13} in multi-agent systems.¹⁰ Because the approach provides a flexible, experiment-agnostic path for building autonomous workflows, it has far-reaching implications for accelerated science across physics, chemistry, materials, and biology.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Phillip Maffettone (pmaffetto@bnl.gov).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All software and data used to generate the results in this manuscript can be found open source under a BSD-3-clause license github.com/matthewcarbone/ScientificValueAgent (tag v1.1.0). Tarballs containing all of the data used to generate our results are hosted open access at <https://doi.org/10.5281/zenodo.8368497>. All figures presented in this manuscript can be regenerated by using these data, and the notebooks stored at the GitHub link above. Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.matt.2023.11.012>.

ACKNOWLEDGMENTS

This research is supported in part by Brookhaven National Laboratory (BNL), Laboratory Directed Research and Development (LDRD) grants no. 22-059, "Precision synthesis of multiscale nanomaterials through AI-guided robotics for advanced catalysts"; no. 23-039, "Extensible robotic beamline scientist for self-driving total scattering studies"; and no. 24-004, "Human-AI-facility integration for the multimodal studies on high-entropy nanoparticles." This research also used resources of the Center for Functional Nanomaterials, the PDF (28-ID-1) Beamline and resources of the National Synchrotron Light Source II, which are US Department of Energy Office of Science User Facilities, at BNL under contract no. DE-SC0012704.

AUTHOR CONTRIBUTIONS

M.R.C. developed the algorithm, led the software development, and performed the *in silico* experiments. H.J.K. and Y.Z. performed the flow reactor experiments. S.Y. provided expertise and guidance in the algorithm development. D.O. performed XRD measurements. D.O., H.J., B.D., B.R., Y.Z., and P.M.M. provided insight on applications at user facilities. P.M.M. conceived the project with M.R.C. and advised the derivation and development. The manuscript was prepared by P.M.M. and M.R.C. and edited by all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 25, 2023

Revised: September 25, 2023

Accepted: November 9, 2023

Published: December 12, 2023

REFERENCES

- Häse, F., Roch, L.M., and Aspuru-Guzik, A. (2019). Next-generation experimentation with self-driving laboratories. *Trends in Chemistry* 1, 282–291.
- Pendleton, I.M., Cattabriga, G., Li, Z., Najeeb, M.A., Friedler, S.A., Norquist, A.J., Chan, E.M., and Schrier, J. (2019). Experiment Specification, Capture and Laboratory Automation Technology (ESCALATE): a software pipeline for automated chemical experimentation and data management. *MRS Commun* 9, 846–859.
- Senior, A.W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Židek, A., Nelson, A.W.R., Bridgland, A., et al. (2020). Improved protein structure prediction using potentials from deep learning. *Nature* 577, 706–710.
- Narayanan, H., Dingfelder, F., Butté, A., Lorenzen, N., Sokolov, M., and Arosio, P. (2021). Machine Learning for Biologics: Opportunities for Protein Engineering, Developability, and Formulation. *Trends Pharmacol. Sci.* 42, 151–165.
- Stach, E., DeCost, B., Kusne, A.G., Hattrick-Simpers, J., Brown, K.A., Reyes, K.G., Schrier, J., Billinge, S., Buonassisi, T., Foster, I., et al. (2021). Autonomous experimentation systems for materials development: A community perspective. *Matter* 4, 2702–2726.
- Campbell, S.I., Allan, D.B., Barbour, A.M., Olds, D., Rakitin, M.S., Smith, R., and Wilkins, S.B. (2021). Outlook for artificial intelligence and machine learning at the NSLS-II. *Mach. Learn. Sci. Technol.* 2, 013001.
- Barbour, A., Campbell, S., Caswell, T., Fukuto, M., Hanwell, M., Kiss, A., Konstantinova, T., Laasch, R., Maffettone, P., Ravel, B., and Olds, D. (2022). Advancing discovery with artificial intelligence and machine learning at NSLS-II. *Sync. Rad. News* 35, 44–50.
- Konstantinova, T., Maffettone, P.M., Ravel, B., Campbell, S.I., Barbour, A.M., and Olds, D. (2022). Machine learning enabling high-throughput and remote operations at large-scale user facilities. *Digital Discovery* 1, 413–426.
- Seifrid, M., Hattrick-Simpers, J., Aspuru-Guzik, A., Kalil, T., and Cranford, S. (2022). Reaching critical MASS: crowdsourcing designs for the next generation of materials acceleration platforms. *Matter* 5, 1972–1976.
- Maffettone, P.M., Allan, D.B., Campbell, S.I., Carbone, M.R., Caswell, T.A., DeCost, B.L., Gavrilov, D., Hanwell, M.D., Joress, H., Lynch, J., et al. (2023). Self-driving multimodal studies at user facilities. In *AI4Mat Workshop 36th Conference on Neural Information Processing Systems*.
- Noack, M.M., Zwart, P.H., Ushizima, D.M., Fukuto, M., Yager, K.G., Elbert, K.C., Murray, C.B., Stein, A., Doerk, G.S., Tsai, E.H.R., et al. (2021). Gaussian processes for autonomous data acquisition at large-scale synchrotron and neutron facilities. *Nat. Rev. Phys.* 3, 685–697.
- Maffettone, P.M., Lynch, J.K., Caswell, T.A., Cook, C.E., Campbell, S.I., and Olds, D. (2021). Gaming the beamlines—employing reinforcement learning to maximize scientific outcomes at large-scale user facilities. *Mach. Learn. Sci. Technol.* 2, 025025.
- McDannald, A., Frontzek, M., Savici, A.T., Doucet, M., Rodriguez, E.E., Meuse, K., Opsahl-Ong, J., Samarov, D., Takeuchi, I., Ratcliff, W., and Kusne, A.G. (2022). On-the-fly autonomous control of neutron diffraction via physics-informed Bayesian active learning. *Appl. Phys. Rev.* 9, 021408.
- McDannald, A., Joress, H., DeCost, B., Baumann, A.E., Kusne, A.G., Choudhary, K., Yildirim, T., Siderius, D.W., Wong-Ng, W., Allen, A.J., et al. (2022). Reproducible sorbent materials foundry for carbon capture at scale. *Cell Rep. Phys. Sci.* 3, 101063.
- Maffettone, P.M., Campbell, S., Hanwell, M.D., Wilkins, S., and Olds, D. (2022). Delivering real-time multi-modal materials analysis with enterprise beamlines. *Cell Rep. Phys. Sci.* 3, 101112.
- Li, J., Li, J., Liu, R., Tu, Y., Li, Y., Cheng, J., He, T., and Zhu, X. (2020). Autonomous discovery of optically active chiral inorganic perovskite nanocrystals through an intelligent cloud lab. *Nat. Commun.* 11, 2046.
- Hua, X., Eggeman, A.S., Castillo-Martínez, E., Robert, R., Geddes, H.S., Lu, Z., Pickard, C.J., Meng, W., Wiaderek, K.M., Pereira, N., et al. (2021). Revisiting metal fluorides as lithium-ion battery cathodes. *Nat. Mater.* 20, 841–850.
- Olds, D. (2020). Synchrotron X-ray Diffraction for Energy and Environmental Materials: The Current Role and Future Directions of Total Scattering Beamlines in the Functional Material Scientific Ecosystem. *Sync. Rad. News* 33, 4–10.
- Kusne, A.G., Yu, H., Wu, C., Zhang, H., Hattrick-Simpers, J., DeCost, B., Sarker, S., Osos, C., Toher, C., Curtarolo, S., et al. (2020). On-the-fly closed-loop materials discovery via Bayesian active learning. *Nat. Commun.* 11, 5966.
- Joress, H., DeCost, B.L., Sarker, S., Braun, T.M., Jilani, S., Smith, R., Ward, L., Laws, K.J., Mehta, A., and Hattrick-Simpers, J.R. (2020). A high-throughput structural and electrochemical study of metallic glass formation in Ni–Ti–Al. *ACS Comb. Sci.* 22, 330–338.
- Caramelli, D., Granda, J.M., Mehr, S.H.M., Cambié, D., Henson, A.B., and Cronin, L. (2021). Discovering New Chemistry with an Autonomous Robotic Platform Driven by a Reactivity-Seeking Neural Network. *ACS Cent. Sci.* 7, 1821–1830.
- Maffettone, P.M., Banko, L., Cui, P., Lysogorskiy, Y., Little, M.A., Olds, D., Ludwig, A., and Cooper, A.I. (2021). Crystallography companion agent for high-throughput materials discovery. *Nat. Comput. Sci.* 1, 290–297.
- Box, G.E.P., Hunter, J.S., and Hunter, W.G. (2005). *Statistics for experimenters*. Wiley Series in Probability and Statistics (Wiley-Blackwell).
- Konstantinova, T., Wiegart, L., Rakitin, M., DeGennaro, A.M., and Barbour, A.M. (2022). Machine Learning for analysis of speckle dynamics: quantification and outlier detection. *Phys. Rev. Res.* 4, 033228.
- Koplik, G., Borggren, N., Voisin, S., Angeloro, G., Hineman, J., Johnson, T., and Bendich, P. (2022). Topological Simplification of Signals for Inference and Approximate Reconstruction. Preprint at arXiv.
- Kusne, A.G., and McDannald, A. (2023). Scalable multi-agent lab framework for lab optimization. *Matter* 6, 1880–1893.
- Balandat, M., Karrer, B., Jiang, D.R., Daulton, S., Letham, B., Wilson, A.G., and Bakshy, E. (2020). BOTORCH: A Framework for Efficient Monte-Carlo Bayesian Optimization. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, pp. 21524–21538.
- Sutton, R.S., and Barto, A.G. (2018). *Reinforcement Learning: An Introduction* (MIT press).
- Baranes, A., and Oudeyer, P.Y. (2013). Active learning of inverse models with intrinsically motivated goal exploration in robots. *Robot. Autonom. Syst.* 61, 49–73.
- Grizou, J., Points, L.J., Sharma, A., and Cronin, L. (2020). A curious formulation robot enables the discovery of a novel protocell behavior. *Sci. Adv.* 6, eaay4237.
- Henson, A.B., Gromski, P.S., and Cronin, L. (2018). Designing algorithms to aid discovery by chemical robots. *ACS Cent. Sci.* 4, 793–804.
- Burda, Y., Edwards, H., Pathak, D., Storkey, A., Darrell, T., and Efros, A.A. (2018). Large-scale study of curiosity-driven learning. Preprint at arXiv.
- Gil, Y., Greaves, M., Hendler, J., and Hirsh, H. (2014). Amplify scientific discovery with artificial intelligence. *Science* 346, 171–172.
- Lehman, J., and Stanley, K.O. (2011). Abandoning objectives: Evolution through the search for novelty alone. *Evol. Comput.* 19, 189–223.
- Eysenbach, B., Gupta, A., Ibarz, J., and Levine, S. (2018). Diversity is all you need: Learning skills without a reward function. Preprint at arXiv.
- Rzhetsky, A., Foster, J.G., Foster, I.T., and Evans, J.A. (2015). Choosing experiments to accelerate collective discovery. *Proc. Natl. Acad. Sci. USA* 112, 14569–14574.
- Swersky, K., Snoek, J., and Adams, R.P. (2013). Multi-task bayesian optimization. *Advances in Neural Processing Systems*, 26. https://proceedings.neurips.cc/paper_files/paper/2013/file/f33ba15effa5c10e873bf3842afb46a6-Paper.pdf.
- Rasmussen, C.E., and Williams, C.K.I. (2006). *Gaussian Processes for Machine Learning* (MIT Press).
- Shahriari, B., Swersky, K., Wang, Z., Adams, R.P., and De Freitas, N. (2016). Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* 104, 148–175.
- Frazier, P.I. (2018). A tutorial on Bayesian optimization. Preprint at arXiv.

41. Mockus, J., Tiesis, V., and Zilinskas, A. (1978). The Application of Bayesian Methods for Seeking the Extremum. *Towards Global Optimization*, 2 (North-Holland Press), pp. 117–129.
42. Srinivas, N., Krause, A., Kakade, S.M., and Seeger, M. (2009). Gaussian process optimization in the bandit setting: No regret and experimental design. Preprint at arXiv.
43. Fortunato, S., Bergstrom, C.T., Börner, K., Evans, J.A., Helbing, D., Milojević, S., Petersen, A.M., Radicchi, F., Sinatra, R., Uzzi, B., et al. (2018). Science of science. *Science* 359, eaao0185.
44. Terayama, K., Sumita, M., Tamura, R., Payne, D.T., Chahal, M.K., Ishihara, S., and Tsuda, K. (2020). Pushing property limits in materials discovery via boundless objective-free exploration. *Chem. Sci.* 11, 5959–5968.
45. Olds, D., Peterson, P.F., Crawford, M.K., Neilson, J.R., Wang, H.W., Whitfield, P.S., and Page, K. (2017). Combinatorial appraisal of transition states for *in situ* pair distribution function analysis. *J. Appl. Crystallogr.* 50, 1744–1753.
46. Maffettone, P.M., Daly, A.C., and Olds, D. (2021). Constrained non-negative matrix factorization enabling real-time insights of *in situ* and high-throughput experiments. *Appl. Phys. Rev.* 8, 041410.