

PNNL-38576

Development of Automated Atom Probe Tomography capability to study the influence of applied voltage and laser power on the final apparent composition of the analyzed specimen

October 2025

SV Lambeets
MA Ziatdinov
M Tong
MG Wirth
AV Liyu
DE Perea



U.S. DEPARTMENT
of **ENERGY**

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062

www.osti.gov
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: <http://www.ntis.gov>

Development of Automated Atom Probe Tomography capability to study the influence of applied voltage and laser power on the final apparent composition of the analyzed specimen

October 2025

SV Lambeets
MA Ziatdinov
M Tong
MG Wirth
AV Liyu
DE Perea

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Abstract

This study presents the development and implementation of an autonomous Bayesian optimization (BO) framework for controlling and optimizing experimental parameters in Atom Probe Tomography (APT). Using commercial silicon needle samples as a benchmark system, we demonstrate that BO can efficiently navigate the complex parameter space of voltage and laser power to achieve target charge state ratios (specifically $\text{Si}^+ / (\text{Si}^+ + \text{Si}^{2+})$) with minimal experimental evaluations. Our implementation integrates Gaussian Process modeling with the CAMECA atom probe control framework, enabling autonomous adjustment of experimental conditions in real-time. Results show that the algorithm successfully converges to target ratios under different scenarios: maintaining a reference ratio, increasing the ratio (favoring Si^{1+}), and decreasing the ratio (favoring Si^{2+}). The system adapts to specimen evolution during analysis, compensating for changes in apex geometry while maintaining optimization targets. This work establishes a proof of concept for AI-driven optimization in APT, addressing the traditional challenges of manual parameter tuning and paving the way for applications to more complex materials where compositional accuracy is critical.

Summary

This study introduces an automated Bayesian optimization system for Atom Probe Tomography that intelligently controls experimental parameters to achieve specific charge state ratios. Using silicon samples as a test case, we demonstrate the algorithm's ability to efficiently navigate the voltage-laser parameter space, rapidly converging to optimal conditions that produce target $\text{Si}^+ / (\text{Si}^+ + \text{Si}^{2+})$ ratios within approximately 10 optimization steps.

The implementation comprises three Python modules that handle machine learning optimization (`ML_lib.py`), hardware communication (`AI_APT_lib.py`), and workflow orchestration (`Maestro_AIAPT.py`). The system successfully addresses several technical challenges inherent to APT analysis, including binary data interpretation, timing ambiguities in data collection, and compensation for specimen evolution during analysis.

Our testing confirms that the algorithm can effectively maintain, increase, or decrease charge state ratios by appropriately adjusting voltage and laser power combinations. In each scenario, the observed parameter adjustments align with theoretical predictions based on field evaporation physics and the Kingham model of post-ionization. Particularly notable is the system's ability to compensate for the changing specimen geometry during analysis by progressively adjusting parameters to maintain consistent field conditions.

The integration of safety features, particularly voltage reduction during computational operations, allows for extended experimental runs while preserving specimen integrity. This automated approach significantly reduces the reliance on operator expertise and trial-and-error parameter tuning, promising improved efficiency, reproducibility, and accessibility in APT experimentation.

This work establishes a foundation for applying AI-driven optimization to more complex materials and different optimization targets, potentially addressing longstanding challenges in accurate compositional analysis of systems prone to preferential evaporation and detection biases.

Acknowledgments

This research was supported by the **Physical and Computational Sciences Directorate Mission Seed**, under the Laboratory Directed Research and Development (LDRD) Program at Pacific Northwest National Laboratory (PNNL). PNNL is a multi-program national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830.

Acronyms and Abbreviations

AI	Artificial Intelligence
APT	Atom Probe Tomography
BO	Bayesian Optimization
GP	Gaussian Processes
LEAP	Local Electrode Atom Probe
ML	Machine Learning
RBF	Radial Basis Function
UCB	Upper Confidence Bound

Contents

Abstract.....	ii
Summary.....	iii
Acknowledgments.....	iv
Acronyms and Abbreviations	v
1.0 Introduction	1
2.0 Results and discussion	3
2.1 Overview of Auto-APT workflow and its component.....	3
2.2 Key commands and their role.....	6
2.3 Challenges in data collection and conversion	6
2.4 Safety and guardrails	8
2.5 Brief description of the laser-assisted field evaporation in atom probe tomography	8
2.6 Results and discussion.....	9
3.0 Future development and perspectives	16
4.0 Conclusion	18
5.0 References.....	19
Appendix A – List of experiments	A.1

Figures

Figure 1: System architecture diagram showing the modular organization of the Ai-APT system	3
Figure 2: Network architecture diagram showing connection between Ai-APT system and the LEAP server	6
Figure 3: Voltage profile showing safety reductions during data processing phases	8
Figure 4: Kingham plot of Si post ionization in presence of electric field	9
Figure 5: Illustration of iso-probability fronts in voltage-laser parameter space and its evolution while the specimen needle is evaporated	9
Figure 6: Hitmap of the Si specimen used in the experiments.....	10
Figure 7: Score plot measuring the differential between the target ratio and the measured ratio over two initial steps and 50 main optimization steps. The initial ratio is 0.36 and the target ratio is 0.36.....	11
Figure 8: Overlay of the BO exploration path across the laser/voltage parameter space in 2+50 steps and the GP mean plot of the parameter space at the 50 th step. Here the algorithm starts from a ratio of 0.36 and targets a ratio of 0.36.....	12
Figure 9: Score plot measuring the differential between the target ratio and the measured ratio over two initial steps and 50 main optimization steps. The initial ratio is 0.36 and the target ratio is 0.65.....	13

- Figure 10: Score plot measuring the differential between the target ratio and the measured ratio over two initial steps and 50 main optimization steps. The initial ratio is 0.23 and the target ratio is 0.15. 13
- Figure 11: Overlay of the BO exploration path across the laser/voltage parameter space in 2+50 steps and the GP mean plot of the parameter space at the 50th step. Here the algorithm starts from a ratio of 0.36 and targets a ratio of 0.65. 14
- Figure 12: Overlay of the BO exploration path across the laser/voltage parameter space in 2+50 steps and the GP mean plot of the parameter space at the 50th step. Here the algorithm starts from a ratio of 0.25 and targets a ratio of 0.15 15

1.0 Introduction

Atom Probe Tomography (APT) is a powerful material characterization technique, enabling three-dimensional atomic-scale imaging and compositional analysis of materials. Its utility has revolutionized fields such as nanotechnology, metallurgy, and semiconductors, where precise knowledge of the composition and the distribution of elements at the nanoscale level is critical to understanding structure-properties relationships¹. Despite its strengths, the accuracy of the compositional data obtained via APT is highly dependent on experimental parameters such as laser power, detection rate, and specimen voltage. These parameters influence critical aspects of the analysis, from field evaporation behavior to ionization probabilities, and directly shape the resulting mass spectrum. An important challenge in APT experiments is identifying optimal parameter combinations that preserve the sample's true composition in the collected data. This challenge is particularly critical for complex materials, such as oxides, where deviations in the detection of certain elements, especially oxygen, introduce significant errors in compositional analysis²⁻⁶. Accurate and systematic parameter optimization is therefore essential to ensure reliable APT data interpretation.

Traditionally, optimizing APT parameters involves trial-and-error approaches, where operators iteratively adjust experimental conditions while assessing data quality and refining their choices. This process is time-consuming, relies heavily on operator expertise, and does not easily lend itself to systematic exploration of the vast experimental parameter space. Automating this optimization process would provide clear benefits, including reduced time and labor, higher reproducibility, and the potential for data-driven discovery of experimental conditions that might be overlooked in manual optimization.

Machine learning (ML) has recently emerged as a powerful approach for automating experimental workflows, with Bayesian optimization (BO) standing out as a technique particularly well-suited for experimental parameter optimization. BO is a probabilistic machine learning framework for black-box optimization problems, where the objective function is expensive or noisy to evaluate, and the relationship between inputs and outputs is not explicitly known. By constructing a surrogate model of the objective function (often using Gaussian Processes, GPs), BO allows for efficient exploration of the experimental parameter space. Its iterative decision-making process balances exploration (investigating unknown regions of the parameter space) with exploitation (refining areas with promising results), enabling rapid convergence to optimal conditions with a minimal number of evaluations. This makes BO a compelling approach to address the challenges of optimizing APT parameters.

In this study, we focus on implementing and testing Bayesian optimization as a framework for automating the control of APT experimental parameters. As a benchmark system, we use commercial silicon (Si) needle samples and target optimization of the ratio of singly charged (Si^+) amongst the total of Si ions collected which includes Si^+ and the doubly charged (Si^{2+}) ions in the mass spectrum. The $\text{Si}^+ / (\text{Si}^+ + \text{Si}^{2+})$ ratio was chosen because it is well-documented to vary as a function of experimental parameters such as laser pulse energy and specimen voltage. Optimizing this ratio enables us to evaluate whether BO can reliably identify experimental conditions that achieve a specific, predetermined target ratio. While Si samples are relatively straightforward to analyze and may not require extensive optimization in routine studies, they present an ideal benchmark for developing and validating the BO framework due to their reproducibility, availability in large quantities, and historical role as standard samples in APT studies.

The primary focus of this work is the implementation of Bayesian optimization in the context of APT and its evaluation using a well-defined and measurable target: the $\text{Si}^+ / (\text{Si}^+ + \text{Si}^{2+})$ ratio. Specifically, we aim to demonstrate that BO can autonomously determine the experimental conditions necessary to achieve the target ratio with minimal experimental evaluations. Success in this benchmark system serves as a foundation for future work on more complex materials, such as oxides or multi-element systems, where deviations in compositional accuracy are a pressing concern. However, this study is deliberately focused on developing and testing the BO framework itself, leaving applications to more complex materials for future work.

This implementation uses the *gpax* Python package for Gaussian Process-based Bayesian optimization and integrates it with the vendor-provided control framework for a CAMECA atom probe instrument. The system is designed to autonomously adjust experimental conditions (laser intensity and voltage), collect data on the resulting $\text{Si}^+ / (\text{Si}^+ + \text{Si}^{2+})$ ratio, and iteratively refine its predictions to converge to the target. This methodology builds on prior efforts from groups such as Felfer et al.⁷, who developed open-source tools for atom probe instrument control, and Gault et al.⁸, who explored machine learning-enhanced analysis of APT data. While existing work has emphasized data analysis and instrument accessibility, our study represents a complementary but distinct contribution: the automation and optimization of APT experimental parameters in real-time.

By demonstrating that BO can successfully optimize APT experimental conditions using the $\text{Si}^+ / (\text{Si}^+ + \text{Si}^{2+})$ ratio as a benchmark, we establish a critical proof of concept for the broader application of AI-driven optimization in APT. While the ultimate goal is to enable the accurate compositional analysis of more complex materials, such as oxides, this work provides the foundation for using BO to systematically and autonomously optimize experimental conditions, reducing reliance on trial-and-error approaches and paving the way for fully automated APT workflows.

This project aims to transform the APT and its injection system into an autonomous platform capable of exploring and control experiments based on specific parameters such as the types of chemical species detected. To achieve this goal, two main limitations must be addressed:

1. Modify the hardware to enable the APT to communicate electronically with computers.
2. Develop a BO-based algorithm capable of communicating and guiding the APT experiment.

The objectives of this project are:

1. Replace the manual control of the parameters of interest.
2. Install a PNNL-control computer capable of modifying experimental parameters.
3. Deploy an algorithm on the PNNL-control computer capable to orchestrate the OAP experiment.

This deployment has been executed on the LEAP4000 system in PNNL as first test phase but is meant to be extended to the other systems present in PNNL (LEAP5000 and LEAP6000).

2.0 Results and discussion

2.1 Overview of Auto-APT workflow and its component

The overall Ai-APT system has been divided in three separate python algorithms:

1. ML_lib.py (Bolib): implements Bayesian optimization algorithms (blue)
2. AI_APT_lib.py (APTlib) Handles communication with the atom probe hardware (orange)
3. Maestro_AIAPT.py (Mstr): Orchestrates the overall workflow and provides user interface (green)

Initialization Phase

The optimization begins with user-defined parameters collected through the main interface (Mstr):

```
python
1 target_ratio = float(input("What target ratio do you want to pursue? "))
2 num_counts = int(input("Enter the number of counts to collect: "))
3 time_limit = int(input("Enter the time limit (s): "))
4 step_limit = int(input("Enter how many ML steps perform before exiting: "))
```

These inputs initialize the ML system's parameter space. Then the Bolib creates an initial set of random sampling points across the parameter space to efficiently explore voltage and laser power combinations.

Data acquisition loop

For each set of parameters, the system:

1. Sets voltage and laser power through CAMECA's communication protocol
2. Collects ion data until reaching either the count threshold or time limit
3. Processes the collected data to calculate charge state ratios

The core of data collection happens in APTlib:

```
python
1 # Request raw data from atom probe
2 msgToSend = "1 LAServer SendData\n"
3 sent = self.readsocket.send(msgToSend.encode())
4 data = self.readsocket.recv(self.MAX_DATA_BUF_SIZE)
```

Safety & Processing phase

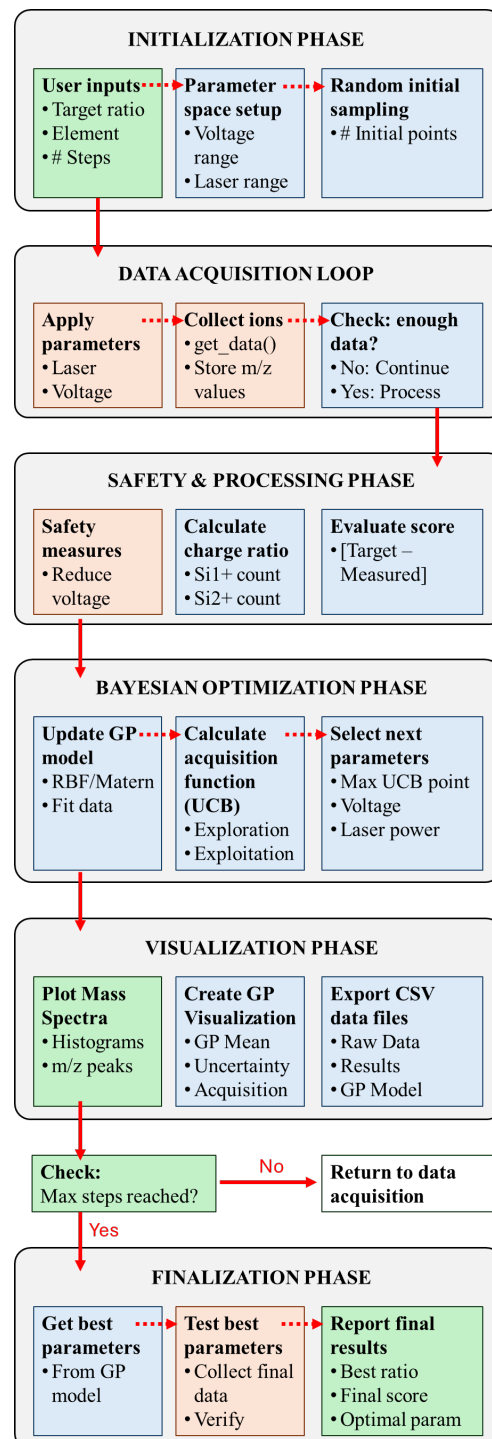


Figure 1: System architecture diagram showing the modular organization of the Ai-APT system

A critical safety feature is the automatic reduction of voltage during computationally intensive operations:

```
python
1 safe_voltage = HvSafe
2 print(f"SAFETY: Reducing voltage to {safe_voltage}V before processing")
3 self._aptlib.set_voltage_log(safe_voltage)
```

The system then processes collected mass spectra to calculate the charge state ratio:

```
python
1 MeasuredRatio = self._MeasureRatio(self._Elements, self._IonArray)
2 Score = abs(self._TargetRatio - MeasuredRatio)
```

This processing identify Si^{1+} ions and Si^{2+} ions and calculates the ratio of $\text{Si}^{1+}/\text{Si}^{n+}_{\text{tot}}$. As first approach we have defined all the ions with a m/z comprised between 27 and 31 Da as Si^{1+} and all the ions between 13 and 16 Da as Si^{2+} .

Bayesian Optimization Phase

The heart of the system is the Gaussian Process (GP) model that learns from each measurement:

```
python
1 # Initialize GP model with chosen kernel
2 if self._KernelType == "Matern":
3     gp_model = gpax.ExactGP(2, kernel='Matern')
4 else:
5     gp_model = gpax.ExactGP(2, kernel='RBF')
6
7 # Fit the model to observed data
8 gp_model.fit(rng_key1, X_norm, self._ScoreArray)
9
10 # Compute acquisition function to determine next point
11 obj = gpax.acquisition.UCB(rng_key2, gp_model, jnp.array(X_unmeasured_norm),
12                          beta=1, maximize=False, noiseless=False)
```

The algorithm objective is to generate a probabilistic recommendation for the optimal set of parameters at each step. Unlike conventional neural networks, which use fixed weights and provide point estimates, Bayesian neural networks (BNNs) treat weights as probability distributions, incorporating uncertainty into their predictions. Due to their probabilistic nature, BNNs are more robust to overfitting, offering a more nuanced understanding of the model's confidence in its predictions, especially when dealing with limited or noisy data. Given that this is a low dimensional but potentially non-linear problem, the BNN architecture will be limited to a simple few-layer perceptron with hyperbolic tangent activations. In comparison of the most commonly used Gaussian process (GP), the BNN is faster at the prediction stage, and scales much better with increasing data size. For the voltage recommendations, we use a standard Bayesian inference approach:

$$p(W, \sigma_{\text{obs}}|Y, X) \propto p(Y|X, W, \sigma_{\text{obs}})p(W)p(\sigma_{\text{obs}}) \quad (1)$$

where X and Y are the parameter sets (voltage and laser power) and composition values, respectively, $p(W, \sigma_{\text{obs}}|Y, X)$ is the posterior distribution representing our beliefs about the BNN weights (W) and the model noise level (σ_{obs}) after seeing the available data (Y, X), $p(Y|X, W, \sigma_{\text{obs}})$ is the likelihood function and $p(W)$ and $p(\sigma_{\text{obs}})$ are the prior distributions over the BNN weights and the noise level. The posterior in (1) can be used to derive a probabilistic recommendation about the next value of applied parameters by performing Markov chain Monte Carlo (MCMC)

sampling from it, $\{W^{(i)}, \sigma_{\text{obs}}^{(i)}\}_{i=1}^N \sim p(W, \sigma_{\text{obs}}|Y, X)$. Then, the next parameter set is selected according to

$$x_{\text{next}} = \operatorname{argmax}_{x_*} (\mu_* + \beta \sigma_*^2) \quad (2)$$

$$\mu_* = \frac{1}{N} \sum_{i=1}^N g(x_*; W^{(i)}, \sigma_{\text{obs}}^{(i)}), \quad (3)$$

$$\sigma_*^2 = \frac{1}{N-1} \sum_{i=1}^N \left(g(x_*; W^{(i)}, \sigma_{\text{obs}}^{(i)}) - \mu_* \right)^2 \quad (4)$$

where g is a Bayesian neural network model, x_* is the space of all possible voltage values, and β is a coefficient controlling the exploration-exploitation trade-off.

The choice between RBF (Radial Basis Function) and Matern kernels significantly impacts the optimization behavior:

RBF Kernel:

- Assumes an infinitely differentiable (extremely smooth) relationship between parameters and charge state ratios
- Produces smoother predictions with strong correlations between nearby points
- More confident in its predictions but potentially less robust to abrupt changes

Matern Kernel:

- Allows for less smooth functions with finite differentiability
- Better captures physical processes that may have transitions or discontinuities
- More conservative in predictions, especially with limited data
- More flexible for modeling realistic physical relationships in the atom probe
- Mathematically expressed as a generalization that includes a smoothness parameter

The model balances exploration of unknown regions with exploitation of promising areas using the Upper Confidence Bound (UCB) acquisition function. The next experimental point is selected to maximize information gain. While we first used the RBF kernel, we quickly used only the Matern kernel because of its better time efficiency.

Visualization Phase

After each measurement, comprehensive visualizations and their corresponding .csv data files are generated including:

1. GP standard deviation maps showing prediction uncertainty
2. Acquisition function plots highlighting promising regions
3. GP mean predictions showing the expected performance across parameter space
4. Mass spectra histograms showing the raw data

Finalization Phase

Upon reaching the step limit, the system identifies the best parameters from the GP model:

```
python
1 best_params = ML_LIB.GetBestParameters()
2 if best_params:
3     best_voltage, best_laser, predicted_score = best_params
```

It then verifies these parameters by collecting a final dataset and reporting the achieved charge state ratio.

2.2 Key commands and their role

The system uses specific commands from the CAMECA Imago LEAP Command Protocol to interact with the atom probe hardware. Every LEAP system follows the same overall architecture: the different controls, gauges and detectors of the instrument are connected to a main server that communicates with a PC_client that handles the user interface. The PNNL computer on which the Ai-APT algorithm is running has been connected to the LEAP4000 server via an ethernet splitter:

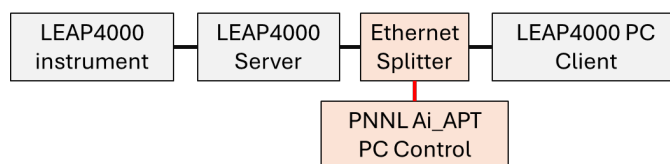


Figure 2: Network architecture diagram showing connection between Ai-APT system and the LEAP server

The connections to the server are established through three different ports:

- The command port 8080 handles the instrument status information and data collection:

```
# Connect to command port (8080)
self.readsocket.connect((self.IP, self.COMMAND_PORT))
```
- The command port 1000 handles voltage control:

```
# Connect to voltage control port (1000)
self.hvcontrol.connect((self.IP, self.CONTROL_PORT))
```
- The command port 1100 handles laser control:

```
# Connect to laser control port (1100)
self.lasercontrol.connect((self.IP, self.LASER_PORT))
```

CAMECA engineers provided the following command to provide the Ai-APT algorithm some controls of the instrument:

- Hardware controls commands:
 - **"1 LASControl ACQ AUTO VOLT ENA FALSE\n"**: Disables automatic voltage control, enabling manual parameter adjustment
 - **"1 LASControl HV SET SPECIMEN VOLTS {volts}\n"**: Sets specimen voltage to specific value
 - **"1 LPS ENERGY SET POINT SET {power} 2\n"**: Sets laser power in picojoules
- Data acquisition commands:
 - **"1 LASServer SendData\n"**: Requests ion detection data from the atom probe
 - **"1 LASServer SendStats\n"**: Requests system statistics including voltage and event rates

These commands form the critical interface between the optimization algorithm and the physical hardware.

2.3 Challenges in data collection and conversion

Binary data interpretation, structure and conversion

The atom probe delivers raw binary data that must be carefully decoded. Each ion detection is represented as a series of 32 bit DWORD called “records”. Each records have many tags, but only four of interest in the context of this work:

Tag 15 (0x0F): X position of the ion, tag 16 (0x10): Y position of the ion, tag 17 (0x11): time of flight of the ion, tag 18 (0x12): mass-to-charge ratio.

The units of the mass record are $1/1024^{\text{th}}$ of one Dalton and the X_POS and Y_POS units are in “LSBs” and are defined by the calibration done at the factory (in our LEAP4000 model, LSBs are in the order of 0.0387mm).

The different conversions are handled as soon as the raw data collected:

```
python
1 # Apply conversion factors
2 v for ion in ions:
3 v     if 15 in ion: # X position
4         ion[15] *= self.X_LSB # 0.03811 mm
5 v     if 16 in ion: # Y position
6         ion[16] *= self.Y_LSB # 0.03815 mm
7 v     if 18 in ion: # Mass-to-charge
8         ion[18] /= self.DALTON_CONVERSION # 1024
```

Data collection timing ambiguity

A significant challenge arises from the atom probe's data collection mechanism:

```
python
1 # Request data that has accumulated since the last request
2 msgToSend = "1 LASServer SendData\n"
3 data = self.readsocket.recv(self.MAX_DATA_BUF_SIZE)
```

Each data request retrieves all ions detected since the previous request, creating potential ambiguity about parameter correlation. If voltage or laser settings were changed recently, the collected dataset may contain ions from multiple parameter configurations. The system must account for this by:

- Implementing appropriate settling time after parameter changes
- Clearing accumulated data before starting a new measurement
- Ensuring sufficient statistics at each parameter setting

Multiple ion detection filtering

When multiple ions are detected from a single laser pulse, data quality can be compromised:

```
python
1 # In the atom probe stats, these are tracked:
2 "dwTotalEvents": event_stats_data[1],
3 "dMultiplePercent": event_stats_data[5],
```

The system explicitly ignores these multiple-hit events to maintain data quality, but this reduces data collection efficiency. This filtering mechanism helps ensure that mass-to-charge calculations remain accurate but requires longer collection times to achieve sufficient statistics.

2.4 Safety and guardrails

APT is a destructive technique. Inherently, the specimen changes over time because APT requires removing material to analyze it, but the constant allowing to extract the surface atoms: the field evaporation remains the same.

During an APT analysis, the applied voltage and laser are used to increase the surface electric field until it starts to extract surface atoms as ions. The value of the electric field at the surface also conditions the charge status of those ions. While evaporating the surface, the apex radius increases, and, subsequently, the intensities of applied voltage and/or laser power also increase to reach the same applied electric field.

One of the main challenges was to implement sufficient guardrails for the algorithm to balance its capability to explore the space of parameters but in the same time, avoid parameter situations that are so destructive for the specimen that it changes it too much. While we are expecting the BO to track a “moving target” it is important to establish enough guard rails.

The solution implemented is to include a systematic voltage decrease of several thousands of volts after each data collection to preserve the sample while the data processing that can take several minutes due to current hardware limitations.

This approach is directly visible in the voltage plot of the experiment where we can directly voltage peaks every time the system is collecting data to make a better decision:

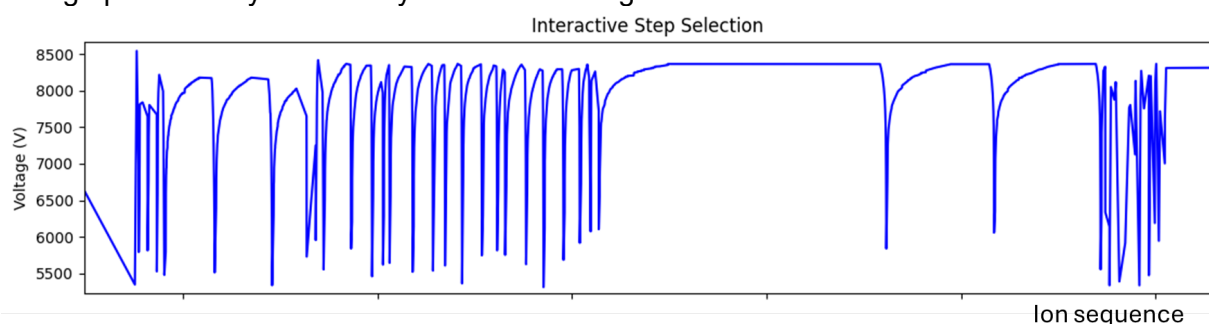


Figure 3: Voltage profile showing safety reductions during data processing phases

2.5 Brief description of the laser-assisted field evaporation in atom probe tomography

Field evaporation is a natural phenomenon in which surface atoms rupture into ions under the influence of a high electric field. This process is probabilistic and is influenced by the quantum tunneling effect and thermal activation. To increase the probability of field evaporation, one can either increase the electric field or raise the temperature of the specimen's apex, as the latter reduces the energy barrier for ionization. In modern atom probe tomography, this dual approach is achieved by combining a high applied voltage with laser-assisted heating.

The applied voltage directly generates a strong electric field at the apex of the needle-shaped specimen, while laser illumination provides thermal energy, briefly increasing the local temperature at the apex. As the specimen is heated, less electric field is required to achieve the same probability of field evaporation. For instance, silicon atoms initially require approximately 33 V/nm at 60 K to field evaporate, but at 300 K, the required electric field would be drastically reduced. This effect is observable in the mass spectra, where the relative abundance of differently charged ions shifts. In voltage mode, the process primarily produces doubly charged ions (Si^{2+}), but under laser-assisted conditions, singly charged ions (Si^{1+}) become more prominent due to the reduced field requirement. A phenomena reported by DR Kingham in his fundamental work focusing on the post-ionization of field evaporated ions, predicting the ratio of different Si^{n+} species depending on the applied electric field on the surface (Figure 4).

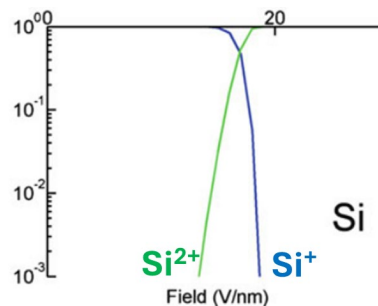


Figure 4: Kingham plot of Si post ionization in presence of electric field

Another critical aspect of field evaporation is the evolving geometry of the needle-shaped specimen during analysis. As atoms are removed layer-by-layer, the apex becomes duller and more voluminous, altering its geometry. Consequently, sustaining a constant electric field—and hence a consistent probability of field evaporation—requires progressively higher applied voltage and potentially greater laser power to maintain adequate thermal energy at the apex.

If we were to map the probability of field evaporation as a function of voltage and laser power, we could visualize "iso-probability fronts" (Figure 5), which represent combinations of voltage and laser power that yield the same probability of evaporation. As the specimen becomes duller during the experiment, these iso-probability fronts shift upward and to the right, reflecting the need for increased voltage and laser power to compensate for the geometric changes and sustain the field evaporation process.

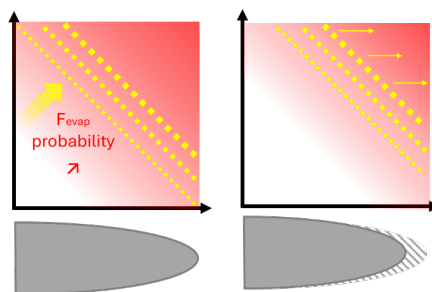


Figure 5: Illustration of iso-probability fronts in voltage-laser parameter space and its evolution while the specimen needle is evaporated

2.6 Results and discussion

While numerous tests were conducted to adjust the safety guardrails and observe the behavior of the algorithm, we will present in this section the three final tests confirming the effectiveness of the algorithm in its current version.

Each test is conducted on the same Si specimen which presents a 001 configuration, easily recognizable by the hitmap of the evaporated atoms on the detector forming a cross. After laser calibration in its X,Y axis and its focus, the sample is run in the APT following default values of analysis. The system is programmed to analyze the Si specimen using a 60pJ laser power pulsing at 125kHz, while adjusting the applied voltage in sort to get an average of 0.5% detection rate. This rate means that, in average, the system collects 625 signals per second.

Launching the Ai-APT system, the automatic control of the voltage and the laser position are suspended, giving up the controls of the voltage and laser power to the Ai-APT system. A first ratio analysis is performed: the system collects by default 2000 counts before returning a ratio. This ratio is collected three times and averaged to be used as a reference ratio r_{ref} from which a target ratio r_{tar} will be determined arbitrarily.

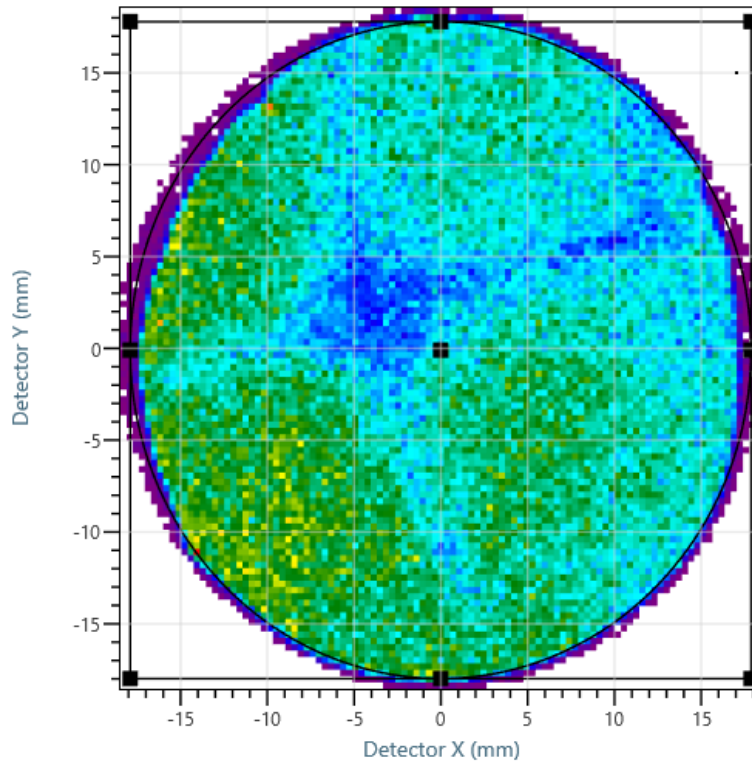


Figure 6: Hitmap of the Si specimen used in the experiments.

All the tests were performed by requesting the collection of 5000 counts over a time period limited to 60 seconds. The number of initial steps were set to 2 while the number of machine learning steps were set to 50.

$r_{tar} = r_{ref}$ case

In this test the Si ratio of interest is manually found is 0.36 and the target ratio is set at 0.36. The Ai-APT algorithm is supposed to seek the same ratio and, logically, we can expect that the system finds an optimal set of parameters close to the original ones.

At the beginning of the experiment, this Si specimen needed 7156V and 60pJ to get a Si ratio of 0.36 and the final optimal result found by Ai-APT is 7356V and 63pJ. When we examine the evolution of the score, which is the difference between the ratio measured and the target ratio, we can see that the algorithm is rapidly finding a promising set of parameters, reaching the optimal zone of parameter by the step 10. The examination of the GP Mean plot and how the parameters are explored, we can see that the algorithm explore the limits of the parameter ranges before rapidly find a zone of interest to focus on.

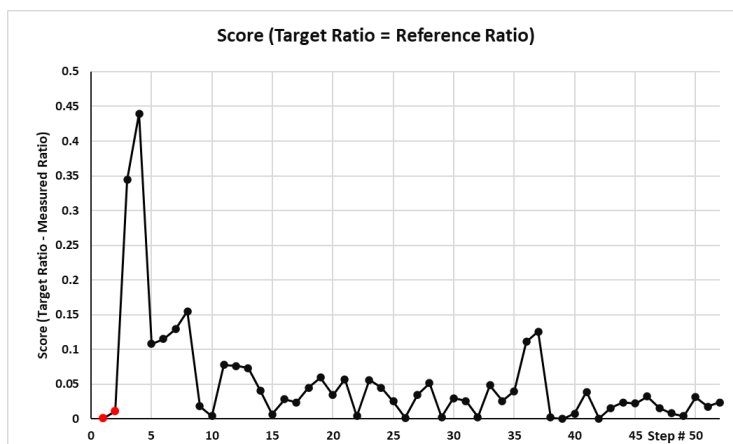


Figure 7: Score plot measuring the differential between the target ratio and the measured ratio over two initial steps and 50 main optimization steps. The initial ratio is 0.36 and the target ratio is 0.36.

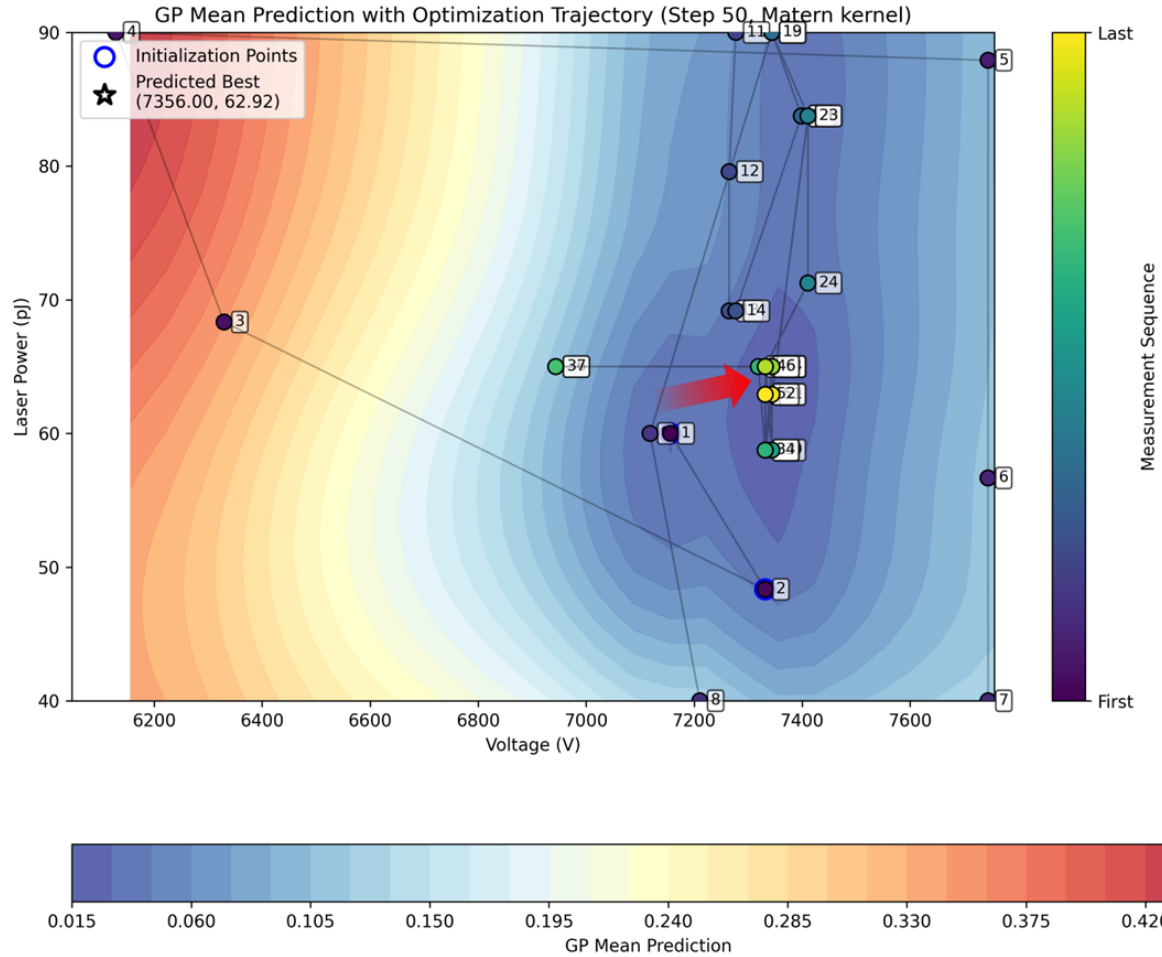


Figure 8: Overlay of the BO exploration path across the laser/voltage parameter space in 2+50 steps and the GP mean plot of the parameter space at the 50th step. Here the algorithm starts from a ratio of 0.36 and targets a ratio of 0.36.

While the final set of parameters found is close to the original parameters, we still can observe a difference with the voltage increasing and the laser increasing. Normally, for a fixed ratio we are looking for a fixed applied electric field and subsequently a fixed set of parameters. This difference is explained by the fact that each measure requires it to probe the specimen, which makes it evolve. After each measure, a quantity of Si atoms is removed from the specimen which increases the radius curvature of the apex. This change of radius curvature means that to reach the same level of electric field, we need to increase the applied voltage.

$r_{tar} > r_{ref}$ case

When we request to increase the ratio, the algorithm will seek a set of parameters that favor the emergence of Si^{1+} vs Si^{2+} . In this particular case, the specimen was presenting a ratio of 0.36 with an applied voltage of 7268 V and a laser power of 60pJ. The target ratio, set at 0.65, is reached by reducing the voltage and increasing the laser power to, respectively, 7018V and 100pJ. On the GP Mean plot, we can observe the behavior of the algorithm that rapidly migrates toward higher laser powers and lower applied voltage by step 6, reaching the upper limits of the laser power.

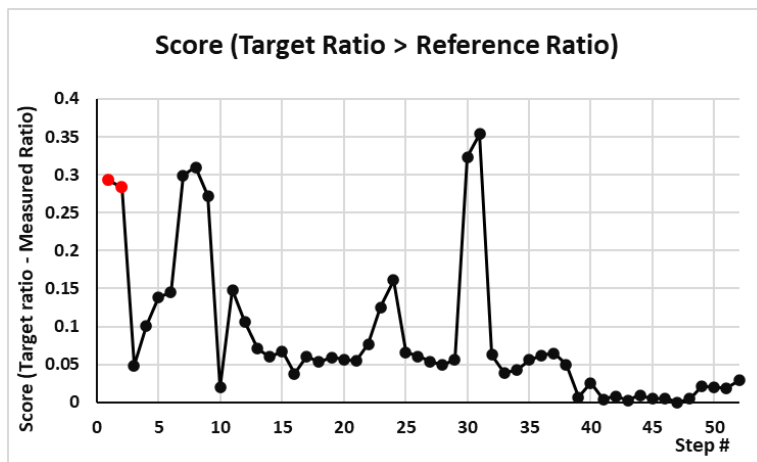


Figure 9: Score plot measuring the differential between the target ratio and the measured ratio over two initial steps and 50 main optimization steps. The initial ratio is 0.36 and the target ratio is 0.65.

Such behavior is expected and follows the trend predicted by DR Kingham plots. By increasing the laser power, the system increases the probability of field evaporation while requiring a lower applied electric field.

 $r_{tar} < r_{ref}$ case

When we request to increase the ratio, the algorithm will seek a set of parameters that favor the emergence of Si^{2+} vs Si^{1+} . In this particular case, the specimen was presenting a ratio of 0.23 with an applied voltage of 8175 V and a laser power of 60pJ. The target ratio, set at 0.15, is reached by increasing the voltage and decreasing the laser power down to, respectively, ~8675V and 40pJ. On the corresponding GP Mean plot, we can observe that the algorithm reaches the optimal zone of parameters after the 5 step, before exploring with different laser power while maintaining high voltages.

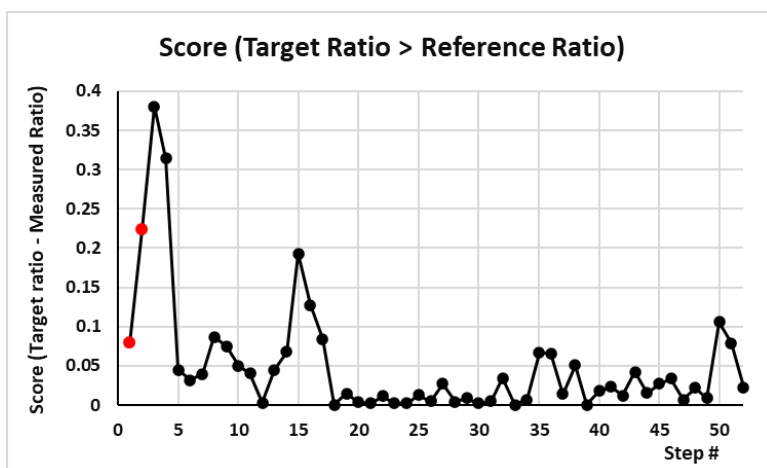


Figure 10: Score plot measuring the differential between the target ratio and the measured ratio over two initial steps and 50 main optimization steps. The initial ratio is 0.23 and the target ratio is 0.15.

As for the other cases this behavior is expected and follows the prediction of DR Kingham. To maintain a reasonable probability of field evaporation while decreasing the laser power, the system relies more on the intensity of the applied electric field. With a more intense electric field during the field evaporation, the occurrence of a second ionization of the Si ions increases as well.

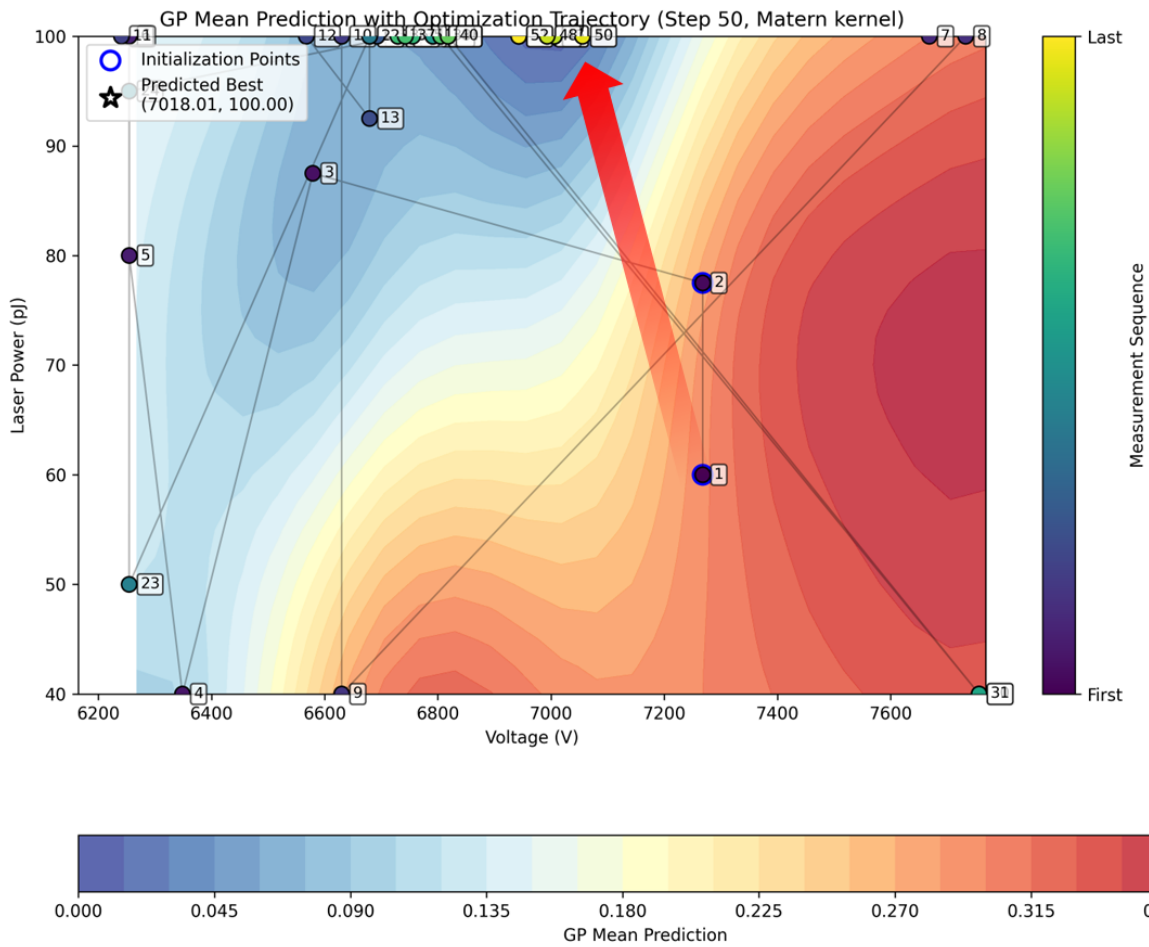


Figure 11: Overlay of the BO exploration path across the laser/voltage parameter space in 2+50 steps and the GP mean plot of the parameter space at the 50th step. Here the algorithm starts from a ratio of 0.36 and targets a ratio of 0.65.

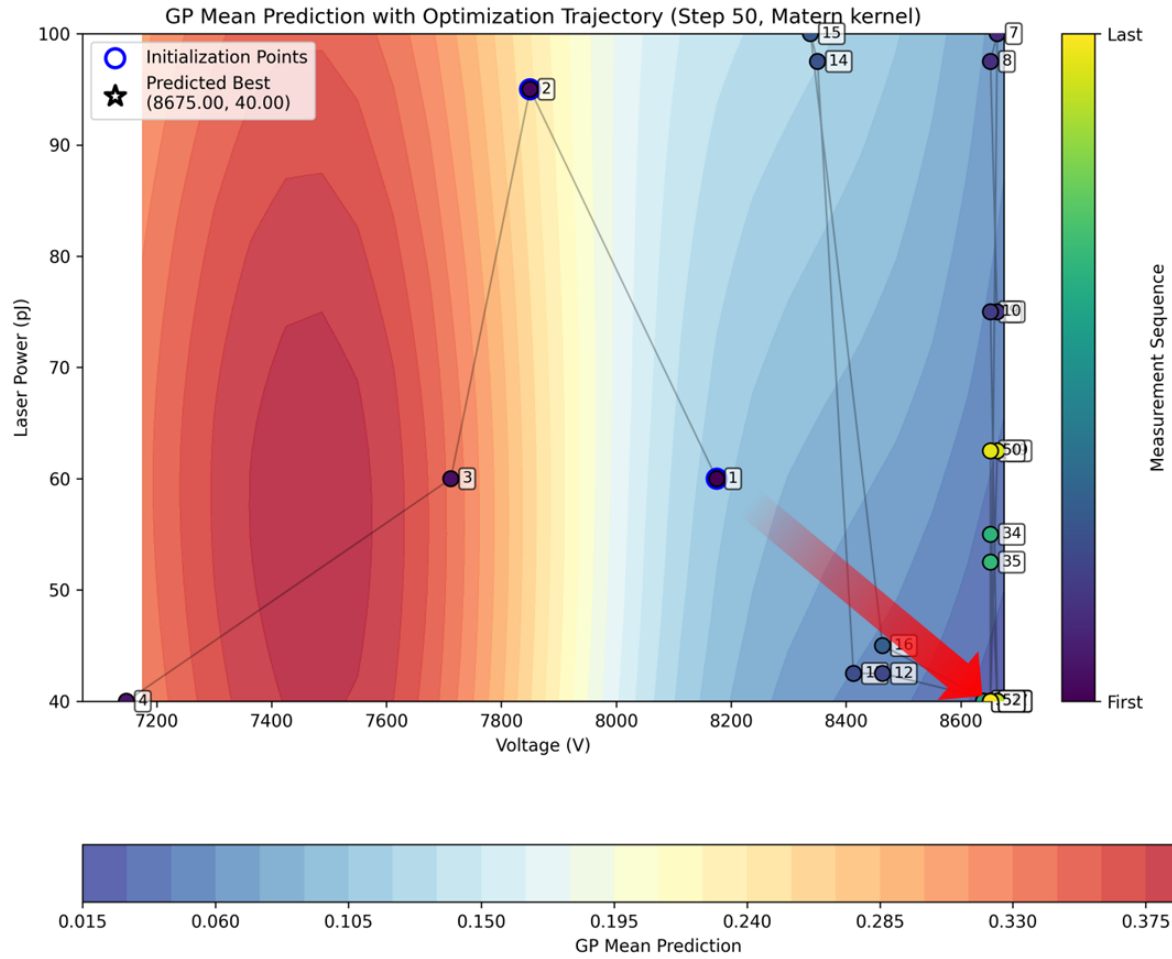


Figure 12: Overlay of the BO exploration path across the laser/voltage parameter space in 2+50 steps and the GP mean plot of the parameter space at the 50th step. Here the algorithm starts from a ratio of 0.25 and targets a ratio of 0.15

3.0 Future development and perspectives

While our implementation successfully demonstrates the viability of BO for controlling voltage and laser power in APT, significant opportunities exist for expanding control capabilities and enhancing safety guardrails. Two distinct paths for expanding system controls emerge: vendor-dependent integration and graphic user-interface (GUI) augmentation.

The vendor-dependent approach involves collaboration with CAMECA to gain access to additional control protocols beyond those currently implemented. This approach offers advantages in reliability and official support, potentially enabling more deeply integrated functionality and access to protected instrument features. However, this path may be limited by vendor prioritization, development timelines, and proprietary restrictions, potentially limiting rapid innovation and customization.

Alternatively, the GUI-based approach involves developing interface automation tools that interact with the existing software interface rather than direct command protocols. This method provides greater flexibility and independence from vendor constraints, allowing for faster implementation of novel features and customized workflows such as combining the instrumental control system to a large language model (LLM). However, this approach may face challenges with robustness, particularly during software updates, and lacks the stability guarantees of vendor-supported solutions. Such customized approach will require long-term in-house support.

Regardless of the integration approach selected, we must acknowledge that the algorithm can control the experiment only as well as its human counterpart would with the same level of controls. The main bottleneck for achieving fully autonomous experiments is providing comprehensive control over all aspects of the experimental process. Pulse frequency modulation represents a critical parameter for optimizing detection efficiency and reducing multiple-hit events. Automated laser positioning control would enable spatial optimization across the specimen apex, potentially improving evaporation uniformity. Similarly, sample positioning optimization could enhance data collection consistency. Detection rate monitoring and automatic adjustment would allow the system to maintain optimal analysis conditions throughout extended runs, particularly important for specimens with varying evaporation characteristics. Moreover, the system requires the capacity to not only identify masses of interest on the mass spectra, but also analyze their relative intensities, profiles, overlays, and other spectral features—capabilities that human operators routinely employ, sometimes passively.

This highlights a fundamental principle: an AI-driven system can only match or exceed human performance when granted the same or greater level of control over experimental parameters. Without equivalent access to all the controls available to human operators, even the most sophisticated algorithm will remain constrained by its limited influence over the experimental apparatus.

A particularly promising development is the integration of electronic valve controls for reactive gas introduction, recently installed and currently in testing phase. This capability would enable automated Field Ion Microscopy (FIM) and Operando Atom Probe (OAP) visualization between APT runs for in-situ sample condition assessment.

Equally important is the development of enhanced safety guardrails that better preserve specimen integrity throughout autonomous operation. Current voltage-reduction measures provide basic protection, but more sophisticated approaches are needed. These include implementing multi-parameter safety envelopes that adapt based on real-time specimen response, detection rate

monitoring with automatic intervention when approaching critical thresholds, and pattern recognition algorithms capable of identifying incipient specimen failure before destructive events occur.

Advanced guardrails could incorporate predictive modeling of specimen evolution during analysis, anticipating geometric changes and preemptively adjusting parameters to maintain stable field evaporation conditions. This would be particularly valuable for heterogeneous materials where evaporation behavior can change dramatically at interfaces or precipitates.

Developing these enhanced controls and safety features while maintaining the core advantages of BO represents the next frontier in autonomous APT experimentation. The ideal implementation would combine the reliability of vendor-approved controls with the flexibility of user-defined automation, creating a system capable of safely exploring complex parameter spaces across diverse material systems while preserving specimen integrity throughout the experimental process.

4.0 Conclusion

The successful implementation and validation of a Bayesian optimization framework for controlling atom probe parameters represents a significant advancement in automating APT experimentation. Our results demonstrate that the BO algorithm can effectively navigate the voltage-laser parameter space to achieve target charge state ratios in silicon samples, converging rapidly within approximately 10 steps across various scenarios. This efficiency highlights the potential for reducing both experiment duration and operator intervention in APT studies. The system's ability to adapt to the evolving specimen geometry during field evaporation is particularly noteworthy. As the specimen apex becomes duller throughout the experiment, the algorithm automatically compensates by adjusting parameters to maintain the desired charge state ratio, demonstrating its robustness in real experimental conditions. This adaptability is crucial for long-duration APT experiments where specimen evolution is inevitable. While this study focused on silicon as a benchmark material with the $\text{Si}^+ / (\text{Si}^+ + \text{Si}^{2+})$ ratio as the optimization target, the framework developed here lays the foundation for extending this approach to more complex materials and different optimization objectives. Future work should explore applications to challenging systems such as oxides, where accurate compositional analysis remains problematic due to preferential evaporation and detection biases. The integration of safety guardrails, particularly the automatic voltage reduction during data processing, ensures specimen preservation during optimization. This feature, combined with the system's ability to handle data collection ambiguities and filter multiple ion detection events, creates a robust framework for reliable APT parameter optimization. In conclusion, this work demonstrates the feasibility and advantages of AI-driven automation in APT experimentation, offering a pathway toward more systematic, reproducible, and efficient materials analysis. By reducing reliance on operator expertise and trial-and-error approaches, this methodology has the potential to accelerate discovery in materials science and expand the accessibility of high-quality APT analysis.

5.0 References

1. Gault, B., Moody, M. P., Cairney, J. M. & Ringer, S. P. *Atom Probe Microscopy*. (2012). doi:10.1007/978-1-4614-3436-8.
2. Cappelli, C., Smart, S., Stowell, H. & Pérez-Huerta, A. Exploring Biases in Atom Probe Tomography Compositional Analysis of Minerals. *Geostand. Geoanalytical Res.* **45**, 457–476 (2021).
3. Hunnestad, K. A. *et al.* Correlating laser energy with compositional and atomic-level information of oxides in atom probe tomography. *Mater. Charact.* **203**, 113085 (2023).
4. Zanuttini, D. *et al.* Simulation of field-induced molecular dissociation in atom-probe tomography: Identification of a neutral emission channel. *Phys. Rev. A* **95**, 1–6 (2017).
5. Gault, B. *et al.* Behavior of molecules and molecular ions near a field emitter. *New J. Phys.* **18**, (2016).
6. Gault, B. *et al.* Influence of the wavelength on the spatial resolution of pulsed-laser atom probe. *J. Appl. Phys.* **110**, (2011).
7. Monajem, M. *et al.* PyCCAPT: A Python Package for Open-Source Atom Probe Instrument Control and Data Calibration. *Microsc. Res. Tech.* 1–12 (2025) doi:10.1002/jemt.70011.
8. Li, Y. *et al.* Machine learning enhanced atom probe tomography analysis. *Prog. Mater. Sci.* **156**, 101561 (2026).

Appendix A – List of experiments

In this section we report the RHIT codes of all the APT runs conducted to test and improve the algorithm. In bold are highlighted the experiments that were completed and used to evaluate the behavior of the algorithm. All raw data are stored in the central APT data server:

\\pnl.gov\Projects\APT_DB\5031.

33520	35550	35551	35558	35595	35596	35683
35684	35691	35697	35704	35705	35707	35708
35709	35710	35715	35780	35782	35842	35845
35979	35980	35982	36203	36204	36220	36223
36306	36309	36310	36317	36319	36320	36321
36322	36323	36324	36325			

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov