



# Data driven drift correction for complex optical systems

Aashwin Mishra, Matthew Seaberg,\* Ryan Roussel, Sanghoon Song, Auralee Edelen, Daniel Ratner and Apurva Mehta

SLAC National Laboratory, 2575 Sand Hill Rd, Menlo Park, CA 94025, USA. \*Correspondence e-mail: seaberg@slac.stanford.edu

Received 16 December 2025

Accepted 31 March 2026

Edited by R. Ingle, University College London, United Kingdom

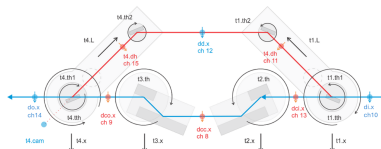
**Keywords:** Bayesian optimization; X-ray free electron laser; X-ray optics; drift correction; split and delay; machine learning.

To exploit the thousand-fold increase in spectral brightness of modern light sources, increasingly intricate experiments are being conducted that demand extremely precise beam trajectory. Maintaining the optimal trajectory over several hours of an experiment with the needed precision necessitates active drift control. Here, we outline time varying Bayesian optimization (TVBO) as a data driven approach for robust drift correction, and illustrate its application for a split and delay optical system composed of six crystals and twelve input dimensions. Using numerical simulations, we exhibit the application of TVBO for linear drift, non-smooth temporal drift as well as constrained TVBO for multi-objective control settings, representing real-life operating conditions. This approach can be easily adapted to other X-ray beam conditioning and guidance systems, including multi-crystal monochromators and grazing-incidence mirrors, to maintain sub-micrometer and nanoradian beam stability over the course of an experiment spanning several hours.

## 1. Introduction

The increasing brightness of light sources, such as the diffraction-limited enhancement of the Advanced Photon Source and the high-repetition-rate enhancement of the Linac Coherent Light Source (LCLS), paves the way for a deeper understanding of fundamental processes at the heart of chemistry, biology and materials sciences. However, these insights necessitate increasingly intricate experiments that demand extreme precision of beam alignment and stability over extended periods (Schoenlein *et al.*, 2016; Margraf *et al.*, 2023). For instance, experiments conducted at LCLS-II-HE will require the X-ray beam to maintain a diameter of just a fraction of a micrometer, with a pointing stability of a few nanoradians at the conclusion of a kilometer-long electron accelerator, a hundred-meter-long undulator section, and greater than one hundred meters of mirror and crystal-based X-ray transport. This configuration needs to be maintained for the entire experiment duration of the order of many hours.

Temporal drifts of beam trajectory can occur in X-ray source points and optical systems due to many factors, such as thermal variations, mechanical vibrations, and environmental changes. Such drift can affect the quality of the data generated. For illustration, due to their short wavelength, X-rays are often used to probe matter at the nanometer scale and as a result X-ray beams must often be focused to sub-micrometer size. Many experiments at the ultra bright X-ray facilities, including picosecond X-ray photon correlation spectroscopy (XPCS) and transient grating spectroscopy, both of which depend on multiple beams maintaining a high degree of overlap on a sample, demand a high degree of beam position



and pointing stability. As such, the beams must have nano-radian level stability on the timescale of the measurement in order to prevent beams losing overlap or a shift to a different region of a heterogeneously evolving sample. This level of stability is typically difficult to meet for a variety of reasons, including thermal/environmental stability and opto-mechanical imperfections. When the required stability cannot be met and when it is not possible to correct for drift parasitically, measurements must be interrupted frequently to correct for errors that compromise data quality, resulting in significant data-collection dead-time and poorer quality measurements.

There are many examples of drift correction techniques and algorithms developed to compensate for time-dependent trajectory alterations, both for lasers and for X-ray sources. These typically rely on a traditional feedback system, using either PID (proportional-integral-derivative) loops or in some cases neural networks (Breitling *et al.*, 2001; Genoud *et al.*, 2011). A critical aspect of such feedback systems is that they must have sufficient diagnostics such that the system can be diagonalized. This can be relatively straightforward for typical laser systems and even for synchrotron X-ray sources (Muller *et al.*, 2012; Martin *et al.*, 2022). In some cases, especially for high-powered lasers or X-ray sources, low-power optical guide beams can be used as a surrogate for the beam of interest (Burkhart *et al.*, 2011; Koehlenbeck *et al.*, 2025). However, in situations where guide beams are not available, and in which the system has more degrees of freedom than the number of independent diagnostic measurements such that traditional feedback cannot be used, alternative approaches must be considered.

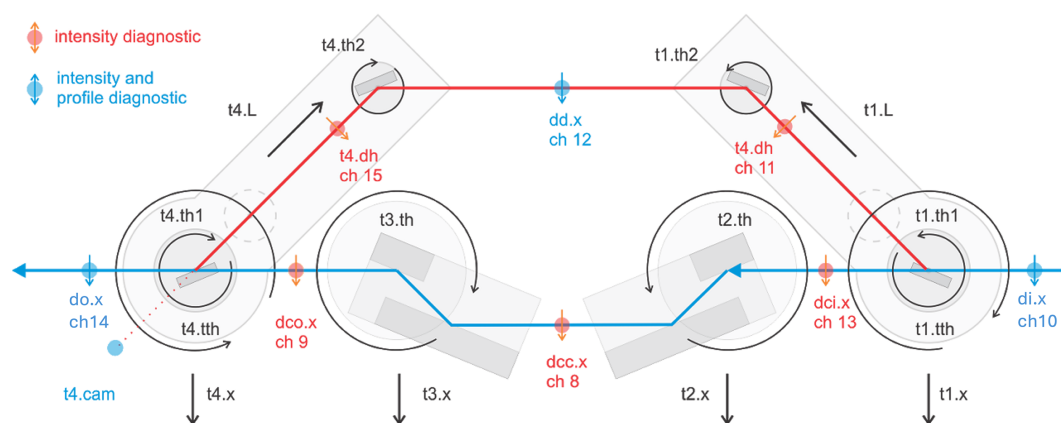
In this investigation, we outline the use of time varying Bayesian optimization (TVBO) as an approach for drift correction in complex optical systems (Kuklev *et al.*, 2022; Kuklev *et al.*, 2023; Xu *et al.*, 2023). We utilize TVBO with a fixed forgetting window approach and apply this for drift correction in the Hard X-Ray Split and Delay system (HXRSND) at LCLS (Zhu *et al.*, 2017). With the introduction of LCLS-II-HE at SLAC, the HXRSND will play a pivotal role in investigations of complex materials using tools for

ultrafast XPCS and transient grating spectroscopy measurements (Gutt *et al.*, 2009). Consequently, it is imperative that we prevent operational inefficiencies from affecting scientific throughput. We report successful application of TVBO for different cases including constant linear drift, non-smooth discontinuous drift and constrained TVBO for multi-objective control settings.

## 2. Methods and application

### 2.1. Time varying Bayesian optimization

Formally, the goal of our drift correction procedure is to optimize a time-varying black-box objective function,  $f_t(\mathbf{x})$ . Here,  $\mathbf{x}$  represent the 12 dimensional motorized degrees of freedom of the HXRSND outlined in Fig. 1. The output of the function is the beam position error, as a function of time  $t$ . As the optical system drifts, the optimal setting in  $\mathbf{x}$  space changes over time. Bayesian optimization (Frazier, 2018; Roussel *et al.*, 2024; Edelen *et al.*, 2019; Mishra *et al.*, 2025) is a sequential sampling approach for finding global optima of black-box functions that are expensive to evaluate, noisy, or have uncertain dynamics, *etc.* It uses a surrogate probabilistic model [often using GPs (Rasmussen & Williams, 2006)] that estimates the distribution of possible function values at points in the domain. In our study, the surrogate model mapping  $\mathbf{x}$  to  $f_t(\mathbf{x})$  is a *Gaussian process* (GP). A GP is a non-parametric model that calculates probability densities over the space of functions, providing a probabilistic surrogate model. While a Gaussian distribution is characterized by a normal distribution with mean and covariance [ $y \simeq \mathcal{N}(\mu, \sigma^2)$ ], a GP is defined via a normal distribution over mean and covariance *functions*,  $y \simeq \text{GP}[m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')] ]$ , where  $m(\mathbf{x})$  and  $k(\mathbf{x}, \mathbf{x}')$  are referred to as the mean and the covariance functions. In this work, we utilize the constant mean function and a Matérn-5/2 kernel to model the covariance kernel, as this kernel is well suited for modeling physical processes that are continuously differentiable. Conditioned on an active dataset,  $\mathcal{D}_t$ , the GP provides a predictive posterior distribution for any untested inputs.



**Figure 1** Schematic of the HXRSND with the channel cut branch in blue and the delay branch in red. The X-ray beam propagates from right to left. The arrows indicate the motorized degrees of the system. In addition, each crystal along the delay branch has ‘chi’ adjustment which corresponds to rotation about the tangential vector of the crystal surface (not shown). The red and blue dots correspond to locations of beam diagnostics, as noted in the legend at the top left of the figure. Figure re-used with permission from Zhu *et al.* (2017).

An acquisition function is used to determine the next point to be sampled at, based on the predictions of the surrogate probabilistic model. The acquisition function balances exploration (*i.e.* a preference towards points where the surrogate probabilistic model’s predictions have high variance) and exploitation (*i.e.* a preference towards points that have better mean predictions according to the model). Common acquisition functions include expected improvement (EI), upper confidence bound (UCB), and probability of improvement (PI) (Garnett, 2023; Snoek *et al.*, 2012). To select the subsequent parameter setting,  $\mathbf{x}^*$ , an acquisition function is optimized over the predictive posterior. As our objective is to minimize the beam position error, we formulate the confidence bound acquisition function as the lower confidence bound (LCB),

$$\alpha_{\text{LCB}}(\mathbf{x}) = \mu_t(\mathbf{x}) - \sqrt{\beta} \sigma_t(\mathbf{x}), \quad (1)$$

where  $\beta$  is often referred to as the exploration parameter. In this acquisition function, the first term  $[\mu_t(\mathbf{x})]$  encourages exploitation by favoring regions predicted to have low error, while the second term  $[\sqrt{\beta} \sigma_t(\mathbf{x})]$  encourages exploration of regions with high uncertainty. The subsequent sample is evaluated at the minimizer of this function,

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}} \alpha_{\text{LCB}}(\mathbf{x}). \quad (2)$$

This sampling is carried out iteratively, where the point is evaluated, the surrogate probabilistic model is retrained with this augmented dataset, and the subsequent point is recommended based on the trained surrogate model’s predictions and the appropriate acquisition strategy.

The key advantage of Bayesian optimization is that it allows optimization in minimal evaluations of the underlying process. This is useful both for time-consuming simulations as well as beam time, where real-time feedback and decision latency are of critical importance. Secondly, it enables us to handle general black box functions without assuming any functional form, via the use of flexible Bayesian models like GPs. Finally, owing to the use of probabilistic surrogate models, Bayesian optimization is robust to noise in the function evaluations (*e.g.* the noise inherent to minimally invasive diagnostics). In this vein, Bayesian optimization has been used for many complex applications such as tuning of particle accelerators (Roussel *et al.*, 2024; Roussel *et al.*, 2023; Edeken *et al.*, 2019).

A limitation of traditional Bayesian optimization is its assumption of the underlying objective function remaining static. TVBO (Brunzema *et al.*, 2022) is an extension that addresses problems where the objective function changes over time. This can occur in many problems, for instance dynamic environments, where the underlying system being optimized exhibits temporal variations due to factors like drift, seasonality, trends, or external influences. The utilization of TVBO for drift correction represents a step beyond classical use of feedback loops for drift correction (like PID controllers) by not just correcting a single variable to a fixed setpoint but also by actively re-optimizing the machine’s performance in an evolving multi-dimensional parameter space as the operating

conditions and the machine itself slowly change. Additionally, the GP surrogate model in TVBO enables robustness to noisy measurements, which are ubiquitous in beamline optics. It learns the underlying function despite stochastic fluctuations.

The central challenge of TVBO is modifying the standard Bayesian optimization framework, which assumes a static objective function, to handle a time varying function. In the recent past, multiple algorithms have been proposed to achieve this, that differ in the manner by which they forget old information, and how they model time dependency. This includes data-centric approaches (Bogunovic *et al.*, 2016; Zhou & Shroff, 2021) like sliding window based TVBO, TVBO with weighing of samples; model-centric approaches (Liu & Djurić, 2021) like utilizing a GP model with a time-dependent kernel, utilizing dynamic GP models; *etc.* In this investigation, we utilize TVBO with a sliding window approach (Zhou & Shroff, 2021) which is now implemented in the Xopt library (Roussel *et al.*, 2023).

Standard Bayesian optimization utilizes the entire history of evaluations,  $\mathcal{D}_{1:t} = \{(x_i, y_i)\}_{i=1}^t$ , to update the surrogate model. However, in a dynamic environment, older samples no longer reflect the current state of the machine. In this study, we employ a sliding window approach of fixed size  $w$ . At any given step  $t$ , the surrogate model is conditioned strictly on the truncated dataset of the most recent  $w$  observations,

$$\mathcal{D}_t = \{(x_i, y_i)\}_{i=t-w}^t. \quad (3)$$

Here, the surrogate model is trained on only the most recent  $w$  samples. This enables the optimization procedure to focus on a recent subset of observations, which may be useful in dynamic environments where significantly older data may no longer be relevant and may even skew the optimization process. The fixed sized window of  $w$  recent samples slides with every subsequent sample. The size of this window controls how quickly the surrogate model can adapt to changes in the objective function.

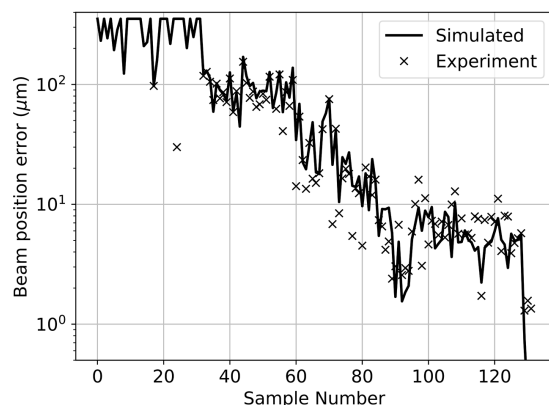
## 2.2. Overview of the HXRSND

In this investigation, we utilize the HXRSND as the system under study for TVBO. As depicted in Fig. 1, this system comprises two branches: the minimally adjustable channel-cut (CC) branch (represented in blue), and the delay branch (illustrated in red), which possesses twelve degrees of freedom used to introduce a delay relative to the CC branch while maintaining a constant trajectory at the system output. The delay range spans from approximately  $-5$  to  $500$  ps, corresponding to the path length difference between the branches ranging from  $-1.5$  to  $150$  mm. The alignment of the HXRSND necessitates a spatial overlap between the two branches at the sample with exceptional precision, as well as optimized intensity at the output. For instance, for many experiments, both branches must be aligned to the same photon energy within approximately  $0.1$  eV with matching intensity, while simultaneously overlapping the resulting beams to a small fraction of the focused beam size.

For measurement techniques such as ultrafast XPCS, the two beams must maintain overlap at the  $\sim 1 \mu\text{m}$  level for many hours, while maintaining a constant path length difference of several millimeters, to maintain data quality (Li *et al.*, 2021). Since the stability of the system is not sufficient for this, the overlap must be manually checked, and if necessary corrected, every  $\sim 10$  minutes, forcing interruption of data collection frequently, requiring a constant operator presence and reducing data-collection efficiency. Typically this manual check involves the invasive insertion of a fluorescent YAG (yttrium aluminium garnet) screen, but in principle potentially non-invasive measurements such as speckle from a static sample can provide the same information (Li *et al.*, 2021). Furthermore, the system is relatively complex from an X-ray optics standpoint, such that, when correction is performed manually, the optical element used for correction may not correspond to the element that caused the drift. Over time, this approach leads to sub-optimal correction and can lead to a need for a more detailed re-alignment of the system.

### 2.3. Details of the simulations

For this study, we performed wave-optical simulations of the system in preparation for the limited availability of XFEL beam time. The simulations are based on decoupled horizontal and vertical ( $2 \times 1\text{D}$ ) wavefront propagation using in-house software (Seaberg, 2025). These simulations model the input beam as fully spatially coherent (a reasonable assumption for XFEL beams) and monochromatic (taking credit for a monochromator upstream of the split and delay system). The various motion degrees of freedom shown in Fig. 1 are all reproduced in the simulation, and care was taken to ensure that the simulation sufficiently captures the dynamics of the operation of the actual HXRSND (see Fig. 2). Small differences between the simulation and experiment may be the result of instability of the XFEL beam into the system, which were not captured by the simulation model. Since the CC branch is intrinsically much more stable than the delay branch, we focus the simulation efforts on maintaining the stability of



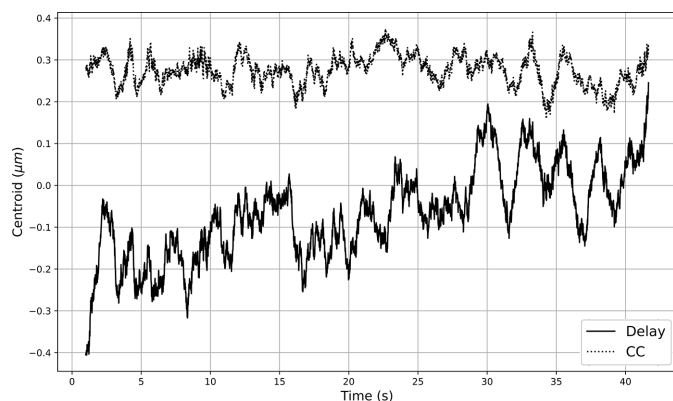
**Figure 2**

Comparison of the predictions of the beam position error from the wave-optical simulations code used in this investigation against experimental results on the HXRSND (delay branch only). Between each sample various degrees of freedom were adjusted both in experiment and also as inputs to the simulator.

the delay branch. In this study, the system was configured for operation at 9.5 keV with zero relative delay between the branches. To judge spatial overlap, we simulate the position of the beam directly at the interaction point as if it were measured using scintillator-based imaging, meaning direct measurement of the beam profile with high signal-to-noise. Even though this represents an invasive measurement, the results of the simulation study also apply to non-invasive measurements that are under development, assuming they provide equivalent information. The non-invasive approach will rely on small-angle speckle measurements, from which sidebands in the autocorrelation provide information about the overlap between the two beams (Sun *et al.*, 2020).

### 2.4. Hyperparameter selection

For the GP model used in this study, the inputs are the settings of the theta and chi knobs of the delay branch crystals, and the output is the predicted beam position error. For the constrained TVBO experiment, an additional GP model accepts inputs of the settings of the theta and chi knobs, and predicts whether the beam intensity for this setting will adhere to the constraint. The parameters of the kernels for the GP models are learned via training using gradient descent. There are other parameters (termed hyperparameters) that are not learned during model training, but have to be selected. These TVBO hyperparameters (specifically, the covariance kernel for the GP model, the acquisition function for the Bayesian optimization, the exploration parameter, the width of the sliding window, *etc.*) were selected based on manual hyperparameter tuning. For the acquisition function, we tested different functions and selected the UCB acquisition function, with an exploration parameter value of  $\beta = 0.1$ . After experiments, we selected a Matern kernel for the GP model, with a smoothness parameter value of  $\nu = 5/2$ . The width of the sliding window is an important parameter controlling the memory of the TVBO procedure. Selecting a window width that is significantly smaller than optimal leads to high variance and noise sensitivity in the optimization, as well as poor uncertainty estimates from the GP model due to a dearth of samples to train on. Contrastingly, a width that is substantially larger than optimal leads to very slow adaption to temporal changes as older, less relevant, data samples dilute the influence of recent and more relevant observations. Additionally, this adds to the computational budget as well, for the GP model has to be trained at every iteration with all samples in the window. After experiments using the simulation, the width of the sliding window was selected as  $w = 40$  samples. In this investigation, the focus is on linear drift observed in the HXRSND; however, for cases with non-linear drift a smaller sliding window width is preferred. The width of the sliding window is reflective of the persistence information content in past samples. In cases with linear drift, a larger sliding window enables better learning of the trend in the drift while minimizing the impact of noise in the samples on this estimation. In corresponding cases with non-linear drift, significantly older samples are less predictive of current behavior and can even



**Figure 3** Visualization of the drift for the horizontal beam position error in the channel cut (CC) and the delay branches observed in an experiment using the HXRSND, without adjustment of any of the degrees of freedom. The rate of drift is observed to be  $\sim 300$  nm in 1 minute, and the standard deviation after subtracting the linear drift is 108 nm r.m.s.

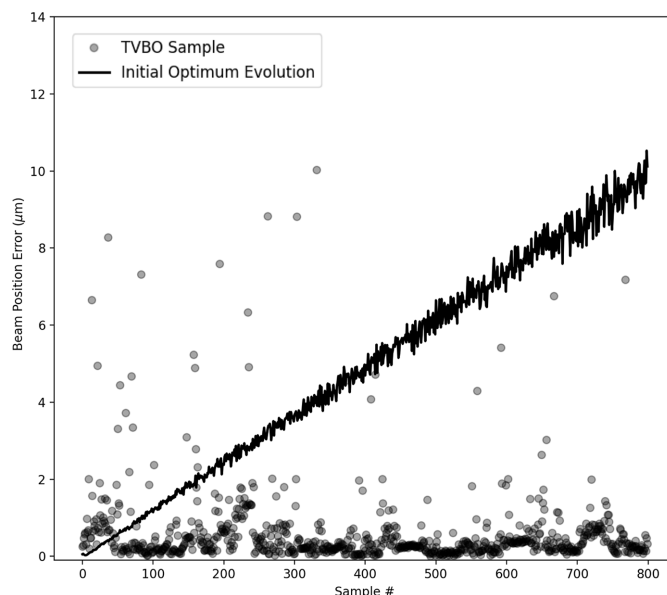
bias the model incorrectly (Zhou & Shroff, 2021), thus a smaller window width is preferred. The rate of drift was estimated from HXRSND operation data, as reported for one experiment in the dataset in Fig. 3, at  $\sim 300$  nm over 60 samples (one minute). The aleatoric noise in the experimental data was calculated to be approximately 100 nm (10% of the stability goal) for the beam position error measurements, and 1% for the beam intensity measurements. These were incorporated in the data during the virtual experiments using the simulations.

### 3. Drift correction for HXRSND operation

In this section, we outline virtual experiments utilizing simulations using TVBO for drift correction under different scenarios such as constant linear drift, drift with discontinuous jumps, and constrained TVBO for multi-objective control settings. The drift is introduced into the angular motion of the crystals, since the system performance relies most heavily upon the stability of these degrees of freedom. For all virtual experiments described here, we assume that samples are taken at a rate of 1 Hz.

#### 3.1. Continuous linear drift

For this scenario, we assume the drift rate constant and equal in magnitude along all input features, along with a smaller stochastic component,  $x_i(t) = x_i(0) + rt + \epsilon$ . The direction of drift (positive or negative) is randomly assigned to each dimension. The goal of this study is to mimic small drifts due to thermal expansion of the system’s constituent components on the minutes to hours timescale. The rate of drift ( $\sim 10 \mu\text{m}$  in 15 minutes) and the variance of the normally distributed stochastic component (100 nm) are inferred from prior experimental data (see Fig. 3). Our objective is to minimize the beam position error, and maintain this minimum over a given period of time as the system drifts. As mentioned in Section 2.2, experiments using the split and delay system typically require overlap checks every  $\sim 10$  minutes to main-



**Figure 4** Results using TVBO for drift correction with a constant, linear drift model, for the beam position error in  $\mu\text{m}$ . The solid line reports the change in performance of the initial optimum setting due to drift. Each semi-transparent circle represents the beam position error for a single sample generated using TVBO. The transparency in the circles is used to assist the reader visualize the location and density of the points even in areas of high overlap.

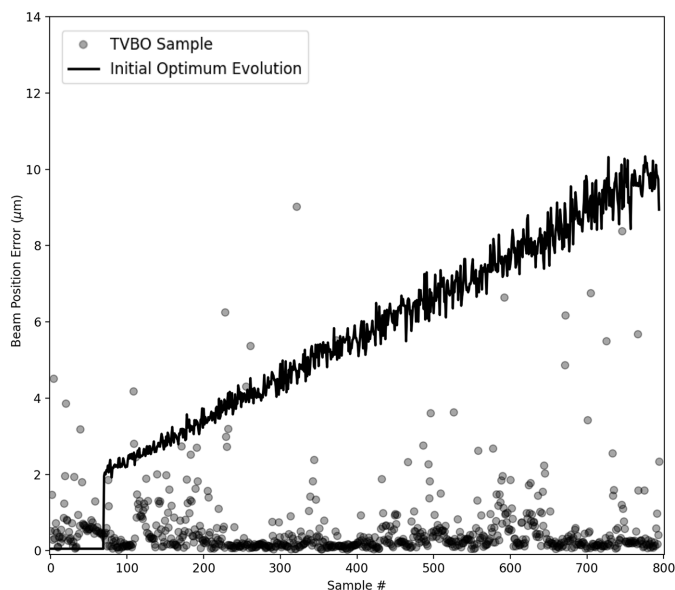
tain  $1 \mu\text{m}$  overlap. As such, the drift rate modeled here is larger than that typically observed in experiments.

The TVBO results are outlined in Fig. 4. In our experiments, 40 initial random samples were generated before commencing the TVBO. In Fig. 4, we show the results after the TVBO procedure is active and is controlling the setting to sample next. The samples generated by the TVBO procedure are reported using dark circles. The solid line reports the evolution of the beam position error in the system if the initial optimum was retained. We observe that the TVBO procedure is able to maintain the beam position error at low values. For the virtual experiment depicted in Fig. 4, the TVBO samples had a mean value of  $0.72 \mu\text{m}$  for the beam position error and a standard deviation of  $1.82 \mu\text{m}$ . This is further bolstered by statistics over 50 runs of the experiment where over 80% of the samples remain within the bound of  $1 \mu\text{m}$  despite the drift.

#### 3.2. Discontinuous drift

In this scenario, we utilize a discontinuous rate of drift, simulating an experiment where drift correction is intermittent due to unavailability of parasitic measurements of the alignment status. During the initial experiment, no data are saved and the drift correction commences after the experiment. This leads to a jump in the drift, as is simulated in this experiment. Consequently, the drift can be mathematically expressed as

$$x_i(t) = \begin{cases} x_i(t) = x_i(0), & \text{if } t \leq T, \\ x_i(t) = x_i(0) + rt + \epsilon, & \text{otherwise.} \end{cases}$$



**Figure 5**

Results using TVBO for drift correction with a discontinuous, episodic drift model. The solid line reports the change in performance of the initial optimum setting due to drift. Each semi-transparent circle represents the beam position error for a single sample generated using TVBO.

Feedback loops for drift correction struggle with such sudden events; for instance, a power supply interruption, that causes non-smooth drift. This affects the robustness of the drift correction.

The results for the experiment using TVBO are outlined in Fig. 5. In the experiments, 70 initial samples were generated before introducing the linear drift with a jump. The samples generated by TVBO are reported using dark circles. The solid line reports the evolution of the beam position error in the system if the initial optimum was retained, where the discontinuity in the value of the beam position error is reported. As is shown in Fig. 4, most of the samples generated by TVBO maintain a low value of the beam position error. For the virtual experiment depicted in Fig. 5, the TVBO samples had a mean value of  $0.88 \mu\text{m}$  for the beam position error and a standard deviation of  $2.04 \mu\text{m}$ . Additionally, over 50 runs of the virtual experiment, over 80% of the samples remain within the bound of  $1 \mu\text{m}$  despite the discontinuous drift.

### 3.3. Drift correction with constraints

In many real-world scenarios, the process of optimization is not solely concerned with minimizing the objective function but to also satisfy inequality constraints on additional outputs, thus defining feasible regions of the solution space. Constrained Bayesian optimization extends standard Bayesian optimization approaches to handle constraint functions, whose form may be unknown, by modeling both the objective and each constraint with separate Gaussian processes. The approach modifies standard acquisition functions to incorporate the probability of constraint satisfaction, using multiplicative terms such as  $P(\text{feasible}|x)$ , enabling the algorithm to balance exploration of feasible regions with exploitation for objective improvement. In the process of drift correction of

the HXRSND, while minimizing the beam position error, we are obligated to maintain the system throughput (*i.e.* beam intensity) at a high value. To this end, we carry out constrained-TVBO experiments, where we minimize the beam position error as the objective while treating the beam intensity value as an inequality constraint, where we maintain the beam intensity at 95% of its initial set value. The setup for these experiments is identical to the linear drift case, but with the addition of the constraint upon the beam intensity.

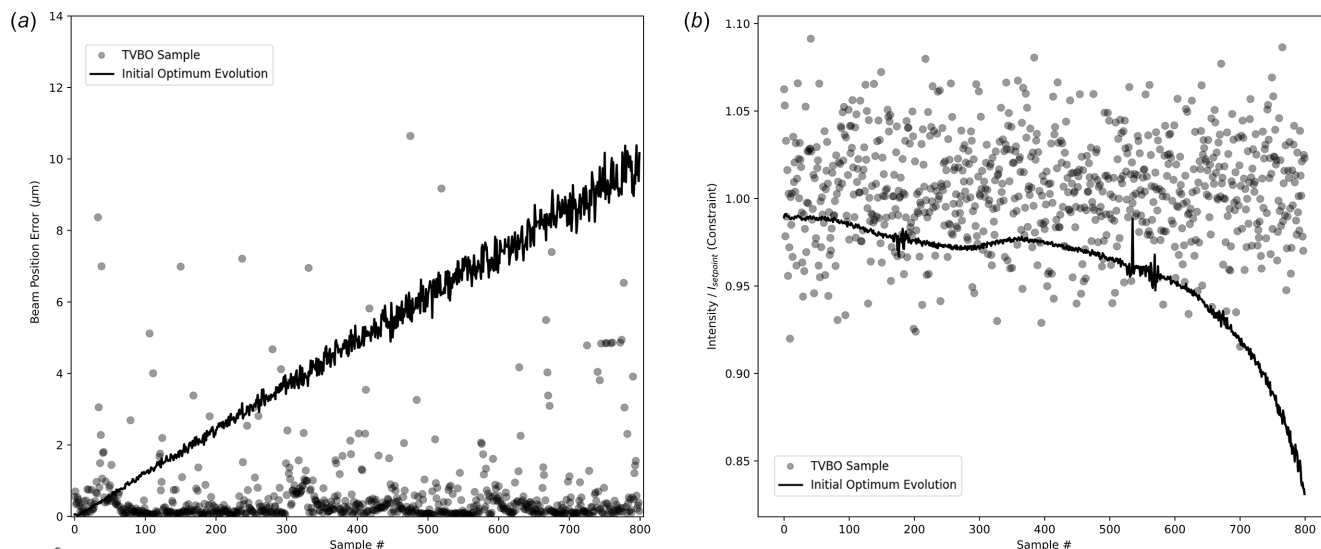
The results for the experiment are outlined in Fig. 6. In Fig. 6(a), the beam position error values using the TVBO procedure are contrasted against the change in the performance of the initial optimum setting. For the virtual experiment depicted in this figure, the TVBO samples had a mean value of  $0.97 \mu\text{m}$  for the beam position error and a standard deviation of  $2.39 \mu\text{m}$ . In Fig. 6(b), the beam intensity (used as the constraint) is visualized and contrasted against the change in the performance of the initial optimum setting due to drift. The contours of the beam intensity are characteristic of crystal optics, characterized by a sharp rise followed by a gently sloping top leading to a sharp drop in value. This is observed in the change in the beam intensity of the initial setting due to drift. It is observed that almost all the samples generated by the TVBO procedure are able to maintain the value of the beam intensity above the specified constraint threshold.

### 3.4. Limitations and deployment challenges

While the wave-optical simulations offer a high-fidelity representation of the beam physics, our study relies on some assumptions within the drift and measurement models. First, the prescribed drift profiles, specifically the continuous linear drift and the idealized discontinuous jumps, serve as tractable proxies for real-world thermal and mechanical instabilities. However, true beamline drift can exhibit more complex, coupled multi-dimensional behavior, including frequency-dependent vibrational noise, thermal breathing and mechanical hysteresis that may not be fully captured by the stochastic noise used in our models.

Furthermore, our measurement model assumes the robust, instantaneous quantification of the beam position error using scintillator-based detection from YAG fluorescence. Put another way, we base the simulated measurements on a direct, high signal-to-noise image of the beam profile at the focus. In practice, our approach relies on the assumption that these results will translate to non-invasive diagnostics currently under development. These real-world diagnostics may suffer from lower signal-to-noise ratios, fluctuating background levels, and intermediate image processing requirements that introduce varying degrees of observational uncertainty.

Transitioning this TVBO framework from a purely virtual environment to live experimental implementation on the beamline presents additional practical challenges that must be navigated. Chief among these are hardware and communication latencies. The current simulation operates sequentially without strictly accounting for the physical delay of motor settling times, network latency within the experimental control



**Figure 6** Results using constrained-TVBO for drift correction. (a) The solid line reports the change in the beam position error of the initial optimum setting due to drift. Each semi-transparent circle represents the beam position error for a single sample generated using TVBO. (b) The solid line reports the change in the beam intensity of the initial optimum setting due to linear drift. Each semi-transparent circle represents the value of the constrained intensity generated by each sample generated using TVBO.  $I_{\text{setpoint}}$  refers to the initial optimum value of the intensity. Thus, the intensity constraint is normalized by the initial optimum value of the intensity. For this virtual experiment using simulations, it was set to  $I/I_{\text{setpoint}} > 0.95$ .

system (e.g. EPICS) and the readout time of diagnostic cameras.

Additionally, there is a computational overhead associated with optimizing the acquisition function and retraining the GP surrogate model at every step. Although the sliding window approach heavily limits the training data size to keep inference times bounded, if the cumulative control loop latency (measurement, inference, and actuation) approaches the characteristic timescale of the drift, the TVBO agent may continuously lag behind the true optimal machine state. Future experimental deployments will require careful synchronization of the TVBO control loop with physical machine latencies and may benefit from incorporating predictive, latency-aware acquisition functions to maintain real-time alignment.

#### 4. Conclusions

The utilization of the increasing brightness of modern light sources opens the opportunity for unprecedented insights into nanoscale and picosecond phenomena through increasingly sophisticated and complex experiments, but for the success of these experiments relies on beam stability with a tight tolerance over extended time periods. The primary hurdle to this end is temporal drift, from myriad and often unknown sources, in the system, that obligates frequent retuning after brief durations, resulting in reduced scientific throughput. In this study, we apply TVBO for temporal drift correction. This technique is applied to a complex optical split and delay system with a high dimensional input space, under different drift models representing real-life scenarios. It is exhibited that TVBO is able to account for and correct temporal drift, resulting in a stable beam. With the advent of stationary Bayesian optimization for beam alignment from a cold start,

TVBO may represent a convenient and proficient approach for drift correction. Using TVBO for drift correction can be adapted to additional X-ray beam conditioning and guidance systems, such as multi-crystal monochromators and grazing-incidence mirrors, to maintain sub-micrometer and nanoradian beam stability over the duration of experiments spanning several hours.

Considering the broader impact of this investigation, transitioning from human-in-the-loop drift correction to TVBO represents a paradigm shift from reactive stabilization to proactive, continuous performance optimization. Current drift correction approaches focus on beam position error only, whereas TVBO continuously seeks an optimum in the presence of drift. In this regard, the beamline would not just be stable, but would attempt to operate at its peak achievable performance at all times, even as thermal, mechanical, and electronic conditions keep changing. Furthermore, TVBO can be used in multi-objective settings using different objective functions and constraints, which is not as simple while using feedback loops. These changes represent a step towards self-driving accelerator facilities, where control transitions from a system of manually tuned, independent feedback loops to an automated agent that can re-optimize over the complex, high-dimensional and time-varying input space.

#### Conflict of interest

The authors declare no conflicts of interest.

#### Data availability

Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

## Funding information

Use of the Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. This work was performed and partially supported by the US Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences Data, Artificial Intelligence and Machine Learning at the DOE Scientific User Facilities program under the MLEExchange Project (award No. 107514).

## References

- Bogunovic, I., Scarlett, J. & Cevher, V. (2016). *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, edited by A. Gretton & C. C. Robert, Vol. 51 of *Proceedings of Machine Learning Research* pp. 314–323. Cadiz, Spain.
- Breitling, F., Weigel, R. S., Downer, M. C. & Tajima, T. (2001). *Rev. Sci. Instrum.* **72**, 1339–1342.
- Brunzema, P., Von Rohr, A. & Trimpe, S. (2022). *61st IEEE Conference on Decision and Control (CDC2022)*, pp. 4046–4052. IEEE.
- Burkhardt, S. C., Bliss, E., Di Nicola, P., Kalantar, D., Lowe-Webb, R., McCarville, T., Nelson, D., Salmon, T., Schindler, T., Villanueva, J. & Wilhelmsen, K. (2011). *Appl. Opt.* **50**, 1136.
- Edelen, A., Neveu, N., Mayes, C., Emma, C. & Ratner, D. (2019). *NeurIPS Machine Learning for the Physical Sciences Workshop (NeurIPS2019)*, Vancouver, Canada.
- Frazier, P. I. (2018). *INFORMS Tutorials in Operations Research*, pp. 255–278.
- Garnett, R. (2023). *Bayesian optimization*. Cambridge University Press.
- Genoud, G., Wojda, F., Burza, M., Persson, A. & Wahlström, C. G. (2011). *Rev. Sci. Instrum.* **82**, 033102.
- Gutt, C., Stadler, L.-M., Duri, A., Autenrieth, T., Leupold, O., Chushkin, Y. & Grübel, G. (2009). *Opt. Express* **17**, 55–61.
- Koehlenbeck, S. M., Lee, L., Balcazar, M. D., Chen, Y., Esposito, V., Hastings, J., Hoffmann, M. C., Huang, Z., Ng, M. L., Price, S., Sato, T., Seaberg, M., Sun, Y., White, A., Zhang, L., Lantz, B. & Zhu, D. (2025). *Light Sci. Applic.* **14**, 129.
- Kuklev, N., Borland, M., Fystro, G., Shang, H. & Sun, Y. (2022). *Proceedings of the 5th North American Particle Accelerator Conference (NAPAC2022)*, pp. 842–845.
- Kuklev, N., Sun, Y., Shang, H., Borland, M. & Fystro, G. (2023). *Proceedings of the 14th International Particle Accelerator Conference (IPAC2023)*, Venice, Italy, *In Proc. 14th International Particle Accelerator Conference*, pp. 4377–4380.
- Li, H., Sun, Y., Vila-Comamala, J., Sato, T., Song, S., Sun, P., Seaberg, M. H., Wang, N., Hastings, J. B., Dunne, M., Fuoss, P., David, C., Sutton, M. & Zhu, D. (2021). *Phys. Rev. Res.* **3**, 043050.
- Liu, Y. & Djurić, P. M. (2021). *Proceedings of the 28th European Signal Processing Conference (EUSIPCO)*, pp. 1462–1466. IEEE.
- Margraf, R., Robles, R., Halavanau, A., Kryzywinski, J., Li, K., MacArthur, J., Osaka, T., Sakdinawat, A., Sato, T., Sun, Y., Tamasaku, K., Huang, Z., Marcus, G. & Zhu, D. (2023). *Nat. Photon.* **17**, 878–882.
- Martin, B. W., Fleischauer, V. E. & Muller, E. (2022). *J. Phys. Conf. Ser.* **2380**, 012088.
- Mishra, A., Seaberg, M., Roussel, R., Poitevin, F., Thayer, J., Ratner, D., Edelen, A. & Mehta, A. (2025). *Synchrotron Radiat. News* **38**(4), 10–17.
- Muller, E. M., Smedley, J., Bohon, J., Yang, X., Gaowei, M., Skinner, J., De Geronimo, G., Sullivan, M., Allaire, M., Keister, J. W., Berman, L. & Héroux, A. (2012). *J. Synchrotron Rad.* **19**, 381–387.
- Rasmussen, C. E., Williams, C. K., *et al.* (2006). *Gaussian processes for machine learning*, Vol. 1. Springer.
- Roussel, R., Edelen, A., Bartnik, A. & Mayes, C. (2023). *Proceedings of the 14th International Particle Accelerator Conference (IPAC2023)*, Venice, Italy, pp. 4847–4850. THPL164.
- Roussel, R., Edelen, A. L., Boltz, T., Kennedy, D., Zhang, Z., Ji, F., Huang, X., Ratner, D., Garcia, A. S., Xu, C., Kaiser, J., Pousa, A. F., Eichler, A., Lübsen, J. O., Isenberg, N. M., Gao, Y., Kuklev, N., Martinez, J., Mustapha, B., Kain, V., Mayes, C., Lin, W., Liuzzo, S. M., St John, J., Streeter, M. J., Lehe, R. & Neiswanger, W. (2024). *Phys. Rev. Accel. Beams* **27**, 084801.
- Schoenlein, R., Adolphsen, C., Mori, R. A., Aquila, A., Bare, S., Bargar, J., Bergmann, U., Boutet, S., Brown, G., Bucksbaum, P., Cargnello, M., Carini, G., Cocco, D., Coffee, R., Cordones-Hahn, A., Cryan, J., Dakovski, G., Devereaux, T., Dunne, M., Durr, H., Emma, P., Evans, I., Fauve, E., Feng, Y., Fletcher, L., Fritz, D., Frost, M., Fry, A., Gaffney, K., Galayda, J., Gleason, A., Glenzer, S., Hastings, J., Hedman, B., Heimann, P., Heinz, T., Hodgson, K., Holtkamp, N., Huang, Z., Kao, C.-C., Kelez, N., Kim, J., Kirchmann, P., Kjaer, K., Lane, T., Lee, J., Lee, W., Levitan, A., Lindenberg, A., Mackinnon, A., Mao, W., Marcus, G., Merminga, L., Minitti, M., Nordlund, D., Ogasawara, H., Perazzo, A., Raubenheimer, T., Reichanadter, M., Reis, D., Robert, A., Robinson, J., Ross, M., Rowen, M., Schlotter, W., Schmerge, J., Schultz, D., Shen, Z.-X., Sokaras, D., Sperling, P., Tavella, F., Thayer, J., Toney, M., Trigo, M., Driel, T. V., Wakatsuki, S., Zhang, L. & Zhu, D. (2016). *LCLS-II High Energy (LCLS-II-HE) a transformative X-ray laser for science*. Technical Report SLAC-R-1143. SLAC National Accelerator Laboratory, Menlo Park, CA, USA.
- Seaberg, M. (2025). *lcls\_beamline\_toolbox*, [https://github.com/mseaberg/lcls\\_beamline\\_toolbox](https://github.com/mseaberg/lcls_beamline_toolbox).
- Snoek, J., Larochelle, H. & Adams, R. P. (2012). *Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS'11)*.
- Sun, Y., Dunne, M., Fuoss, P., Robert, A., Zhu, D., Osaka, T., Yabashi, M. & Sutton, M. (2020). *Phys. Rev. Res.* **2**, 023099.
- Xu, C., Boltz, T., Mochihashi, A., Santamaria Garcia, A., Schuh, M. & Müller, A. S. (2023). *Phys. Rev. Accel. Beams* **26**, 034601.
- Zhou, X. & Shroff, N. (2021). *55th Annual Conference on Information Sciences and Systems (CISS)*, 24–26 March 2021, Baltimore, MD, USA, pp. 1–6. IEEE.
- Zhu, D., Sun, Y., Schafer, D. W., Shi, H., James, J. H., Gumerlock, K. L., Osier, T. O., Whitney, R., Zhang, L., Nicolas, J., Smith, B., Barada, A. H. & Robert, A. (2017). *Proc. SPIE* **10237**, 102370R.