# Machine-learning-guided design of electroanalytical pulse waveforms

Cameron S. Movassaghi,[1,2†*+] Katie A. Perrotta,[1†] Maya E. Curry,[3] Audrey N. Nashner,[1] Katherine K. Nguyen,[1] Mila E. Wesely,[4,5] Miguel Alcañiz Fillol,[6] Chong Liu,[1] Aaron S. Meyer,[7] and Anne M. Andrews[1,2,7,8*]

[1]*Department of Chemistry & Biochemistry, University of California, Los Angeles, Los Angeles, CA 90095, USA*

[2]*California NanoSystems Institute, University of California, Los Angeles, Los Angeles, CA 90095, USA*

[3]*Institute of Society and Genetics, University of California, Los Angeles, Los Angeles, CA 90095, USA*

[4]*Department of Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles, CA 90095, USA*

[5]*Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095, USA*

[6]*Interuniversity Research Institute for Molecular Recognition and Technological Development, Universitat Politècnica de València -Universitat de València, Camino de Vera s/n, Valencia, 46022, Spain*

[7]*Department of Bioengineering, University of California, Los Angeles, Los Angeles, CA 90095, USA*

[8]*Departments of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, and Hatos Center for Neuropharmacology, University of California, Los Angeles, Los Angeles, CA 90095, USA*

---

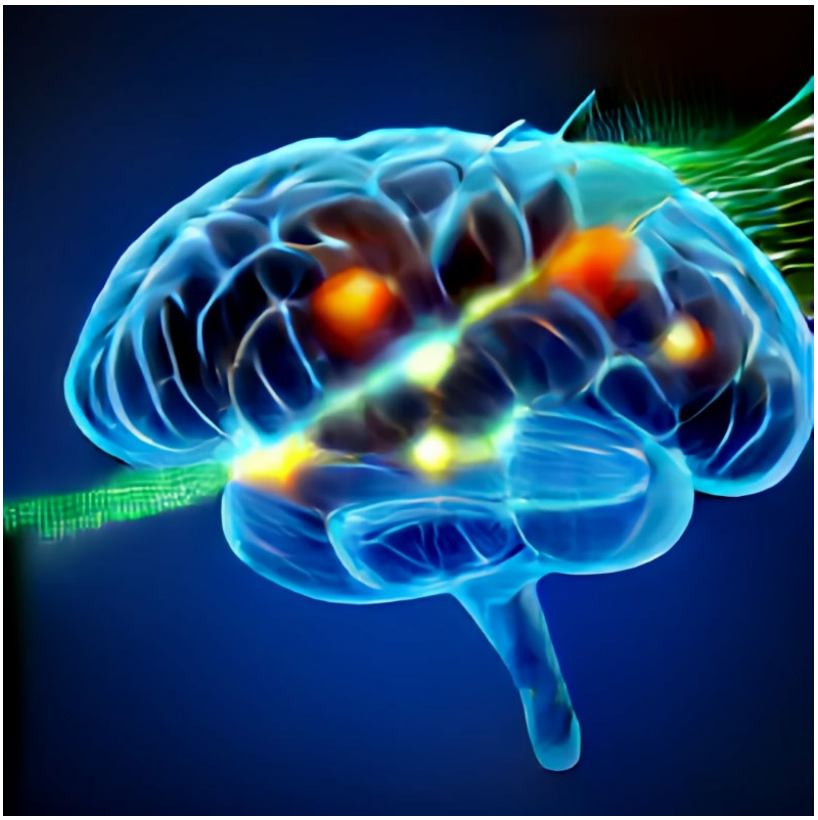
[†]These authors contributed equally to this work

*To whom correspondence should be addressed: aandrews@mednet.ucla.edu and csmova@g.ucla.edu.

[+]Present address: Department of Computational Biomedicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048

**KEYWORDS:** Bayesian optimization, electrochemistry, voltammetry, serotonin, neurotransmitter, brain, *in vivo*

# Graphical Abstract/Cover Art Submission



***About the image:*** Generated and used with permission from Craiyon.com, formerly DALL·E mini, an artificial intelligence (AI) text-to-image generator. The resulting image was returned when the AI was prompted to generate the image of an optimal waveform for serotonin detection.

## Abstract

Voltammetry is widely used to detect and quantify oxidizable or reducible species in complex environments. The neurotransmitter serotonin epitomizes an analyte that is challenging to detect *in situ* due to low concentrations and co-existing similarly structured analytes and interferents. We developed rapid-pulse voltammetry for brain neurotransmitter monitoring due to the high information content elicited from voltage pulses. Generally, the design of voltammetry waveforms remains challenging due to prohibitively large combinatorial search spaces and a lack of design principles. Here, we illustrate how Bayesian optimization can be used to hone searches for optimized rapid pulse waveforms. Our machine-learning-guided workflow (SeroOpt) outperformed random and human-guided waveform designs and is tunable *a priori* to enable selective analyte detection. We interpreted the black box optimizer and found that the logic of machine-learning-guided waveform design reflected domain knowledge. Our approach is straightforward and generalizable for all single and multi-analyte problems requiring optimized electrochemical waveform solutions. Overall, SeroOpt enables data-driven exploration of the waveform design space and a new paradigm in electroanalytical method development.

# Introduction

Voltammetry is widely employed across fields, including energy storage,[1,2] catalysis,[3] materials,[4] organic synthesis,[5] protein engineering,[6,7] and electroanalysis (*i.e.,* neuroscience,[8-12] diagnostics,[13] environmental applications,[14] and food and beverage analysis[15]). Despite the many types of analytes suitable for voltammetry, few design principles exist to enable voltammetry waveforms to be identified and optimized systematically. This lack of objectively guided waveform design and optimization imposes significant limitations on the accuracy, selectivity, and robustness of voltammetry applications for single- or multi-analyte detection and monitoring.

A grand challenge in chemical neuroscience is to uncover the functional and dysfunctional interplay between neurotransmitters in the brain.[16] Voltammetry is broadly used to characterize and quantify electroactive neurotransmitter release and reuptake using brain-implanted electrodes during biological perturbation,[17-19] including in humans.[10] Recent progress has focused on developing novel electrode materials, coatings, or data analysis procedures to improve the selectivity and sensitivity of real-time neurochemical monitoring in behaving subjects.[17,20-27] Meanwhile, voltammetry waveform development (*i.e.,* selecting optimal waveform parameters for detecting a particular analyte) has remained essentially unchanged for decades. It relies principally upon historic performers (*e.g.,* pre-patterned waveforms), heuristics, and grid searches.[28-33]

For neurochemistry applications, historic performers include fast-scan cyclic voltammetry (FSCV) triangle or N-shape waveforms for detecting evoked dopamine[12] or serotonin,[34] respectively, *in vivo*. Modifying these waveforms impacts sensitivity, selectivity, and temporal resolution.[28,35-37] For example, the development of fast-cyclic square-wave voltammetry has improved the sensitivity and selectivity of dopamine[38] and serotonin[39] detection by superimposing triangle and N-shape waveforms, respectively, on pre-patterned staircase waveforms. Other waveform modifications have led to fast-scan controlled absorption voltammetry and multiple cyclic square-wave voltammetry to determine basal dopamine[40] or serotonin levels.[41,42] These approaches required separate waveforms to measure different analytes and timescales and were derived from the prior triangle and N-shape waveforms in a guess-and-check manner (**Fig. 1**, top).

To enable multi-analyte monitoring (*e.g.*, simultaneous serotonin and dopamine detection) across timescales (*i.e.*, quantification of basal and stimulated neurotransmitter levels using the same waveform in the same recording session), we developed rapid pulse voltammetry (RPV).[43] The latter utilizes background-inclusive (*i.e.,* non-background subtracted) data, requiring novel waveform design to produce informative background currents.[44] This custom design is opposed to other popular pulse voltammetry approaches (*e.g.,* normal, differential, staircase), which use pre-patterned approaches over longer time courses (s to min).[45] While also based on characteristic oxidation and reduction potentials derived from the triangle and N-shape waveforms, rapid pulses (*i.e.*, 2 ms), rather than fast linear sweeps, reduced

fouling and produced informative faradaic *and* non-faradaic currents. The resulting current-time fingerprints from our original generation (OG) RPV waveform[43] yielded analyte-specific information that can be used by partial least squares regression (PLSR) or other supervised regression models (*e.g.*, artificial neural networks, elastic net) to distinguish analytes and predict their concentrations. Because the OG waveform was inspired by heuristics from the voltammetric electronic tongue (VET) field for 'soft' sensing (*e.g.,* intermediate and counter pulses),[46-48] we refer to this as VET-inspired design (**Fig. 1**, middle).

Having shown that our VET-inspired OG waveform outperformed conventional waveforms,[43] we sought a generalizable and expandable approach to designing and optimizing rapid pulse (and other types of) waveforms. Because tuning specific waveform parameters improves analyte-specific currents,[17,28,49] we hypothesized that enhanced RPV waveforms for serotonin and dopamine co-detection (and many more analytes) exist but remain undiscovered due to the lack of design principles needed to explore intractable large waveform search spaces.

We focused first on detecting serotonin to address this waveform space problem (*vide infra*). Serotonin is involved in modulating mood, anxiety, and reward-related behavior *via* interconnecting brain circuits.[50-54] Serotonin is also an essential gut hormone and it plays a role in spinal pain transmission and immune function.[55-58] Serotonin is a challenging target to detect using voltammetry due to its relatively low physiological concentrations (high pM to low nM),[51] colocalization with other neurotransmitters having similar redox profiles (*e.g.*, dopamine), and irreversible

oxidation byproducts[59] that can foul electrodes. We further hypothesized that a waveform development paradigm to discover optimized serotonin waveforms would generalize to other neurochemicals and other types of analytes.

When developing RPV or other complex waveforms, a prohibitively large number of waveform step or segment combinations impedes exhaustive empirical investigation, even for a few steps or segments. Step/segment potentials, lengths, order, and hold times are all variables for investigation when exploring and improving waveforms; minor modifications of each variable can have complex effects on electrochemical signals.[35] While a 'guess and check' approach has yielded the handful of useful conventional and VET-inspired waveforms mentioned above, one-parameter-at-a-time or randomized[60,61] optimization approaches do not take advantage of the rich information diversity encoded in complex waveforms, leaving the overall waveform search space relatively unexplored.

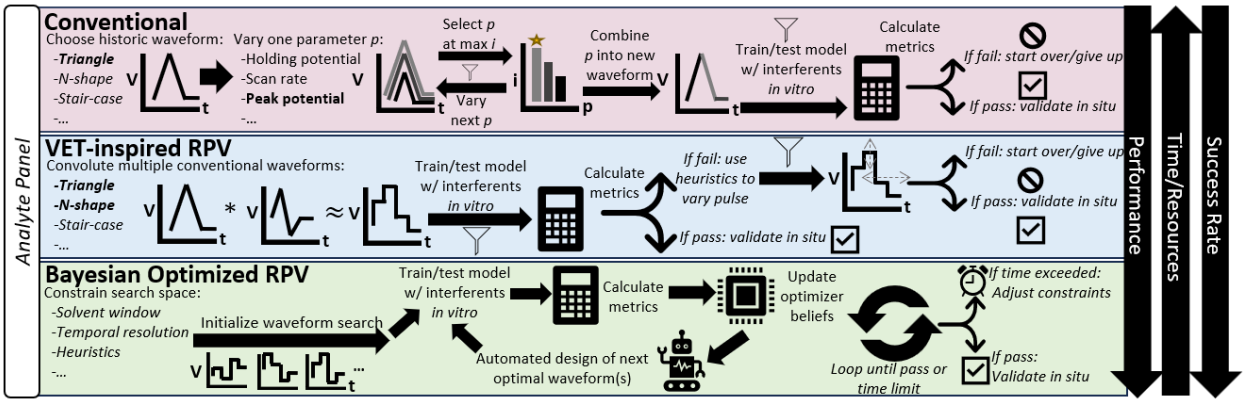



**Figure 1.** Approaches to voltammetry waveform design. Funnels denote likely bottlenecks.

Recently, Bayesian optimization has been used to navigate intractable physiochemical search spaces when combined with experimental training data.[4,7,62] This type of adaptive experimental approach presents an opportunity to pair machine learning with electroanalysis to create a new waveform development paradigm

(**Fig. 1**, bottom). Here, we present a Bayesian optimization workflow (SeroOpt) that generates fit-for-purpose voltammetry waveforms for selective serotonin detection. To our knowledge, a systematic machine-learning-based approach to designing, testing, and optimizing analyte-specific waveforms has not yet been reported. We show that analyte-specific waveform information depends on specific potentials occurring in a particular order and timing, confirming the need for a parsimonious search approach across parameter dimensions. Our active learning approach outperformed randomly designed and domain expert-designed waveforms after only a handful of iterations. Our methods can be straightforwardly extended to designing any voltammetry waveform for any electroactive analyte to discover new and perhaps non-intuitive waveforms optimized for application-specific metrics. To encourage widespread adoption, we provide data, tutorial code notebooks, and videos at https://github.com/csmova/SeroOpt, as well as our corresponding open-source voltammetry acquisition and analysis software at https://github.com/csmova/SeroWare.

# Results

## The SeroOpt workflow casts waveform development as black-box optimization.

We designed the following Bayesian optimization workflow for robust, iterative, and adaptive voltammetry waveform development (SeroOpt; **Fig. 2**). Representative *i-t* curves (*i.e.,* voltammograms) are provided (**Fig. S1**). We sought to identify an input (a rapid pulse waveform) related to an optimal output objective (sensor performance metric; *e.g.*, serotonin detection accuracy) by an unknown, ground-truth objective function (the black box). This function can only be accessed by obtaining experimental training data on various waveform-metric combinations,

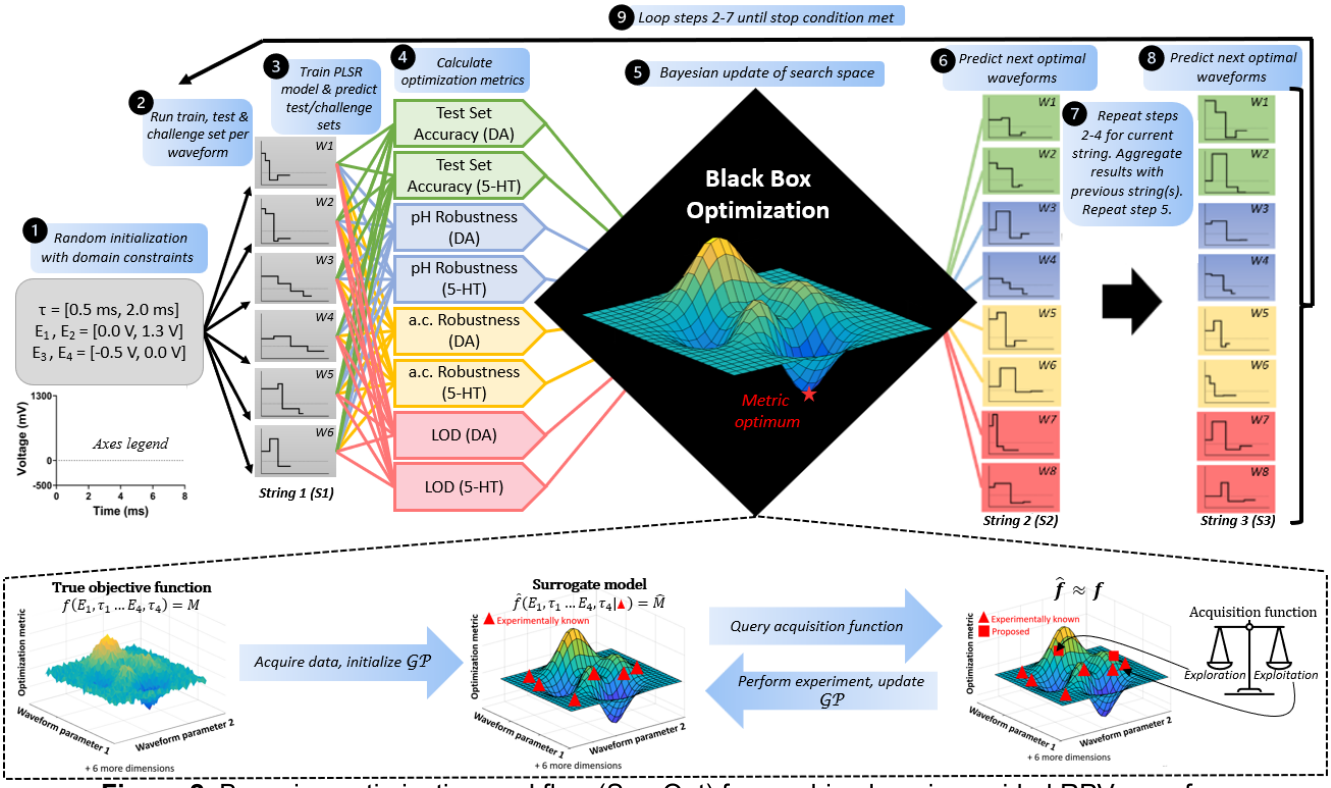



**Figure 2.** Bayesian optimization workflow (SeroOpt) for machine learning-guided RPV waveform design for serotonin (5-HT) and dopamine (DA). An example visualization of optimization landscapes is shown (bottom). GP = Gaussian process, M = metric, W = Waveform, S = String, a.c. = altered cation; ˆ represents estimation of true value

approximating the black-box function using a surrogate model, and then querying the model to generate an input (waveform) corresponding to a predicted objective optimum. The generated (new) waveform is then tested experimentally, and the true objective value for that waveform is provided as subsequent training data for the next round of optimization. When a probabilistic surrogate model is used, both the model predictions (mean) and associated uncertainty (variance) can be updated using Bayesian inference as new data (evidence) becomes available in each iteration. This optimize-update process repeats sequentially, referred to herein as Bayesian optimization. Each of the workflow steps is described in detail below.

### Search space constraints & initialization by embedding domain knowledge.

Each training waveform **W** was embedded as a vector in 8-dimensional space such that $\mathbf{W} \coloneqq [E_1, \tau_1, E_2, \tau_2, E_3, \tau_3, E_4, \tau_4]$ (**Fig. 2,** step 1). Here, $E_i$ is each potential step (V) and $\tau_i$ is each step hold time (ms). In this initial design, for eventual comparison with our original generation (OG) human-designed four-step waveform[43] (**Fig. 3a**), we constrained the search space to four steps per waveform, with $E_1$ and $E_2$ constrained to 0-1.3 V and $E_3$ and $E_4$ constrained to -0.5-0 V. These constraints ensured that waveforms remained inside the solvent window[35] and encoded a 'pulse/counter-pulse' concept (*i.e.,* anodic steps followed by cathodic steps) from VET theory.[63] We constrained $\tau$ to 0.5-2.0 ms based on our preliminary results showing that capacitive current completely decays after ~2 ms, yet critical features are contained in as little as the first ~0.5 ms of each pulse.[43] Pulses do not result in voltage cross-talk (*i.e.*, residual capacitive current from successive voltage steps).[39,40]

The hold time was defined as $(100 - \sum_{i=1}^{4} \tau_i)$ ms to limit the number of parameters. Each pulse was applied at 10 Hz; the holding potential was defined as $E_4$.

To initialize a model of the relationship between waveform and objective (*i.e.*, the optimization metric), six waveforms were randomly generated using the constraints above (**Fig. 2**, step 1). The choice of six waveforms was arbitrary and within the number of waveforms that could be experimentally evaluated in a single-day experiment. We refer to this collection of random initialization waveforms as string 1 (S1).

### Model calibration & optimization metrics allow for relevant objective functions.

We obtained experimental calibration curves (**Table 1**) for each S1 waveform (gray boxes, **Fig. 2**) to train a partial-least squares regression (PLSR) model as demonstrated previously.[43] The PLSR model predicted the test and challenge set sample concentrations of serotonin and dopamine (**Fig. 2**, steps 2-3; see *Methods* for definitions of training, testing, and challenge samples). These predictions were used to calculate the eight optimization metrics listed (**Fig. 2**, step 4; defined in **Table S1**). All metrics were calculated on all waveforms in each string, unless otherwise noted (**Fig. 2**, steps 2-4). We focus on the results for the second waveform (W2) of each string, which is optimized across strings for the serotonin test set prediction accuracy metric. The latter is the mean absolute error in the PLSR model predictions of test samples T1-4 (including a blank; **Table S1**), thus creating a minimization task (maximum accuracy implies minimal error). We chose mean absolute error rather than relative error due to the presence of the blank (null true concentration).

The choice of test set accuracy as an optimization metric was motivated by several factors. First, we pursued single-objective optimization for simplicity and (at the time of analysis) a lack of user-friendly open-source software to perform multi-objective human-in-the-loop optimization. Having to choose only a single metric to focus on, test set accuracy is an attractive choice as it is a direct measure of waveform performance instead of alternatives, such as PLSR model-specific metrics (*e.g.,* scores clustering). The use of model-specific metrics is less physically meaningful and would limit the extendibility of our method. Using physically meaningful parameters such as test set accuracy, our workflow remains model-agnostic (*i.e.,* any model that performs supervised regression prediction can be used). For similar reasons of retaining metrics in raw form, we chose not to combine multiple metrics into a single objective task (*e.g.,* scalarization[64]).

Second, we encoded selectivity in our test and challenge set design. Our calibration curve varies the concentrations of all analytes and interferents across the training, test, and challenge sets used to build and evaluate the PLSR models (**Table 1**). If the PLSR model for a given waveform confuses any interferent for serotonin, this will be represented in the test or challenge set accuracy metric for serotonin as it will contribute to the mean absolute error. Thus, serotonin test and challenge set accuracy is a proxy for selectivity in varying dopamine, 5-HIAA, ascorbate, DOPAC, pH, and $K^+$/$Na^+$ concentrations (see *Methods*).

Lastly, other analytical figures of merit that could be used as optimization metrics (sensitivity, limit of detection (LOD), linear range, *etc.*) are irrelevant if model

accuracy and selectivity are not first established. For example, we included LOD as an alternative optimization metric (**Fig. 2**). The selectivity performance of LOD-optimized waveforms (inferred *via* test and challenge set accuracy) was poor. Thus, we did not continue to optimize for LOD in subsequent campaigns but were still able to utilize these waveforms as training data by calculating their other metrics. For these reasons, we focused on test set accuracy. Specifically, we focused on serotonin (5-HT) because it is historically a more difficult analyte to detect by voltammetry. Its concentrations are approximately 10-fold lower than dopamine in striatum[51] and serotonin has complex redox mechanisms and fouling processes.[34]

Regardless, we included other optimization metrics in our workflow rather than solely serotonin test set accuracy to explore which metrics have an objective landscape that is 'optimizable'. As this was a first attempt, we had no guarantee that the serotonin test set accuracy was a viable choice of metric. We also wanted to investigate other analytes and metrics for future use with multi-objective optimization. For example, we included dopamine-specific metrics in the scheme for comparison with our original RPV work[43] because serotonin/dopamine co-detection is a long-term goal for multi-objective optimization.[65]

To maximize the training data produced in an experimental day, we calculated the performance of all waveforms on all metrics in each string, regardless of which metric a waveform was designed to optimize. For example, the optimal serotonin test set accuracy waveform (W2) in each string was used to calculate the serotonin test set accuracy metric. Still, the performance of this waveform on the dopamine, pH,

and altered cation (a.c.) accuracy metrics was also recorded. This approach allows additional waveforms (albeit waveforms not optimized specifically for that metric) to be tested per string rather than solely the one 'optimized' waveform for each metric. Performing single objective optimization in this parallel manner explores 'optimizable' metrics while obtaining additional training data per string in a simple yet sample-efficient manner. For example, if test set accuracy failed as an optimizable metric for serotonin, we could pivot to an alternative metric exhibiting promising optimization progress (*e.g.*, serotonin pH or a.c. accuracy, or serotonin LOD), with training data already aggregated across all waveforms for that metric.

### Parallel single objective optimization of multiple metrics.

The waveform embeddings and corresponding experimentally determined metrics were used to train the surrogate models (*i.e.*, Gaussian processes)[66] of the unknown objective functions (**Fig. 2**, step 5). As mentioned, only single-objective optimization was performed on each metric. Separate Gaussian processes were

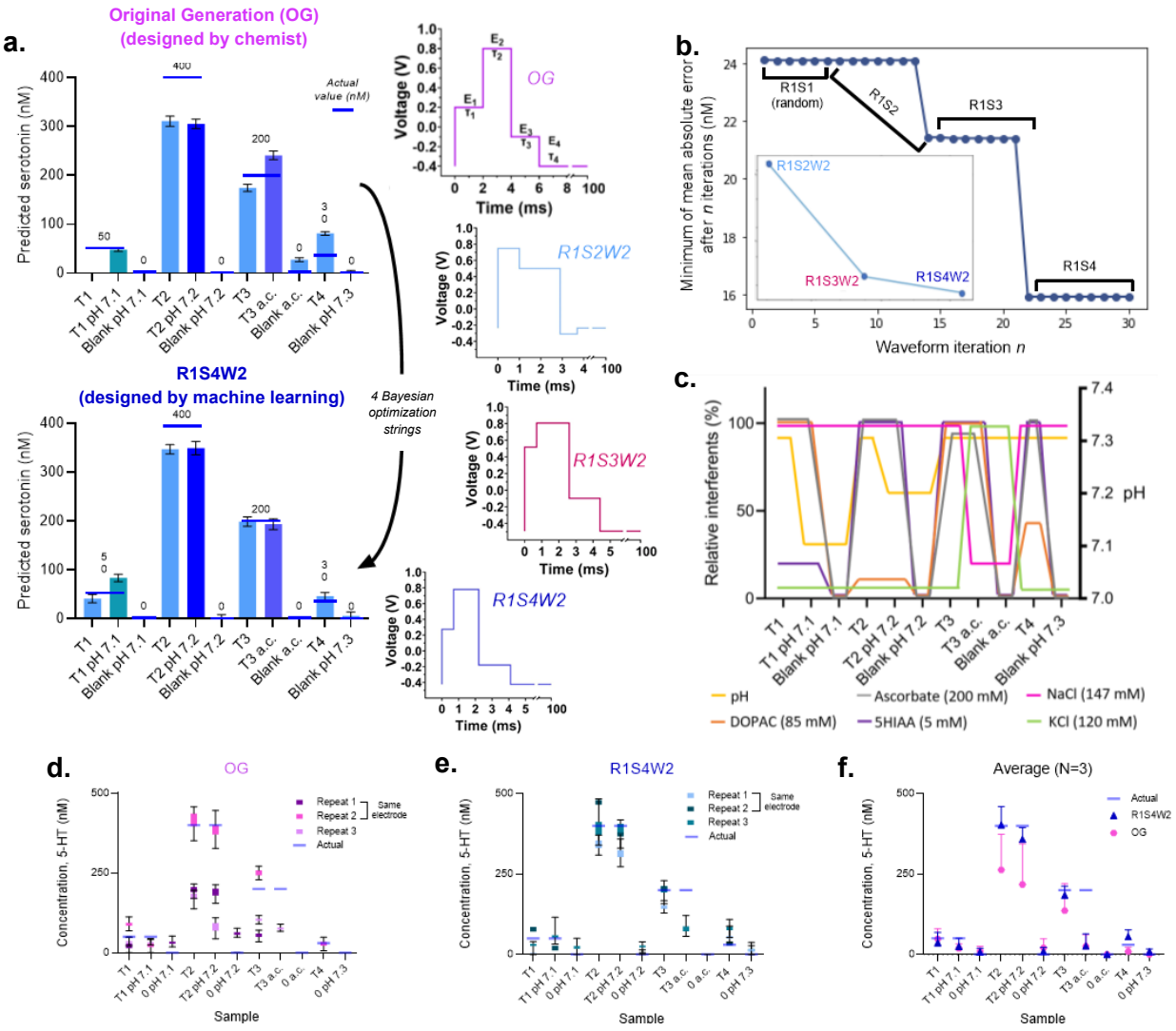



**Figure 3. a)** Bayesian optimization waveform (R1S4W2; bottom) outperforms the original generation (OG) human-designed RPV waveform (top) after four iterations. Error bars represent standard deviation. **b)** Convergence plot of the minima of 5-HT test set accuracy per string. The waveforms optimized specifically for 5-HT test set accuracy (W2) are shown on the inset. **c)** Varied interferents encountered in the test and challenge set samples (a.c. = altered cationic salt concentrations). **d)** Test and challenge set results for the OG waveform in triplicate across two electrodes. Error bars represent minimum and maximum values predicted. **e)** Test and challenge set results for the optimized serotonin waveform (R1S4W2) in triplicate across two electrodes. Error bars represent minimum and maximum values predicted. **f)** Average of d,e. Error bars represent standard deviation.

trained (one for each metric; eight total) in parallel on the aggregated data after evaluating each string. An acquisition function (*i.e.*, expected improvement)[66] finds the optima of each surrogate function and outputs the next most likely waveform that

will improve each respective metric (**Fig. 2**, step 6). The process then repeats (**Fig. 2**, steps 7-9). The overall workflow is illustrated in **Figures 2** and **S2.**

The eight waveforms (each corresponding to optimization for one of the eight metrics) output from the first optimization loop of this workflow are shown as string 2 (S2). The eight new waveforms were generated with each new waveform optimized on a single metric (*i.e.,* using the training data generated from S1 (**Fig. 2**, steps 4-6)). Because S1 was randomly generated to initialize the surrogate model, S2 represented the first iteration of optimized waveforms produced by the workflow.

We repeated the optimization loop by obtaining experimental calibration curve data using each new S2 waveform. We then calculated the individual optimization metrics, aggregated the data with the previous string(s) (*e.g.*, all S3 waveforms were predicted using all S1 *and* S2 data, one metric at a time), and predicted the next set of optimal S3 waveforms for each metric (**Fig. 2,** steps 7-8). This process was repeated again to generate four waveform strings in total (**Fig. 2**, step 9). We refer to the group of strings as S1-4. Each string had eight waveforms (W1-8) corresponding to the eight separate metrics, except the initial string (S1), which had only six randomly generated waveforms (arbitrary). All four strings and their associated waveforms were collectively referred to as run 1 (R1).

### Machine learning outperforms human-guided waveform design.

Across R1, three new waveforms were generated optimized for serotonin test set accuracy (S1W2 was random; the three successive waveforms (S2W2, S3W2,

S4W2) were each more highly optimized than the last. The evolution of the serotonin accuracy waveform across three successive strings was compared to our initial RPV OG waveform (**Fig. 3a**). Similarities were noted between the OG waveform and the final optimized waveform, R1S4W2 (**Fig. 3a**, top right and bottom right, respectively). In the first run, the final waveform generated by our Bayesian optimization scheme nearly perfectly mimicked our chemically intuitive choices for the potentials of the waveform design; the step potentials differed only by ~100 mV or less. The more remarkable differences were in the individually optimized step lengths (τ) for R1S4W2. Values of τ are rarely optimized individually and instead are set to a global value decided by one-factor-at-a-time optimization under single experimental conditions (*e.g.,* 2 ms for all steps in the OG design).[38-41,43]

Even though R1S4W2 was only 5.5 ms long, it outperformed the OG which was 8 ms. Given the similarity in pulse potentials, the increase in data fidelity was attributed partially to changes in the hold times of each step; that is, Bayesian optimization was able to generate better-performing choices of τ.

While a 2.5 ms difference in overall pulse length was ostensibly negligible at data rates of 1 MHz, this equates to a reduction of 2500 data points per scan. This reduction can easily save gigabytes of data that otherwise would have to be stored and save computation time wasted during multi-hour experiments. Decreasing the overall length of the rapid pulse sequence also opens opportunities to increase the temporal resolution to >10 Hz or design more complex pulses with additional steps while retaining 10 Hz sampling.

We did not simply 'get lucky' or stumble across a similar waveform randomly, as the convergence plot (**Fig. 3b**) shows that for each optimization string (S2-S4), the waveform optimized for serotonin test set accuracy (W2) found a new minimum for serotonin prediction error during each iteration. This improvement across strings suggests that the surrogate model is learning a reasonable representation of the optimization landscape for serotonin accuracy. Convergence plots for all metrics and runs are provided (**Fig. S3**).

### Explicit and implicit discovery of interferent-agnostic waveforms.

Next, we compared the results for the test and challenge set samples from the OG waveform to R1S4W2 (which should be the best-yet waveform for test set serotonin performance). Indeed, R1S4W2 outperformed the OG waveform for serotonin detection in the test and challenge set samples (see *Methods*). The train and test sets contain samples with varying levels of three physiologically relevant metabolites (DOPAC, 5-HIAA, ascorbate). Meanwhile, the challenge set samples have physiologically relevant pH, $Na^+$, and $K^+$ levels held constant in the training set (**Fig. 3a,c**, samples denoted pH 7.1, pH 7.2, and altered cations or "a.c."). The optimized serotonin waveform R1S4W2 outperformed the OG waveform for interferents it was explicitly (DOPAC, 5-HIAA, ascorbate) and not expressly (pH, $Na^+/K^+$) trained on.

While the OG waveform confounded changes in pH and $Na^+/K^+$ occurring in the challenge set, the R1S4W2 waveform did not suffer similar pitfalls (see samples T2 pH 7.2, T3 a.c., blank a.c. for each waveform in **Fig. 3a**). We further discuss the

performance of test and challenge set samples in **Fig. S4a,b**. This result was not due to the waveform not sensing a change in current for varying cation concentrations or being 'electrochemically silent'.[67] Increases in current (hundreds of nA) were evident when aCSF a.c. blanks were injected compared to normal aCSF blanks (**Fig. S4c**. Similar responses were noted for pH blanks.

To investigate whether the initial results for R1S4W2 outperforming the OG waveform were precise and robust, the waveforms and training/test/challenge sets were run in triplicate using two different electrodes (**Fig. 3d-f**). We determined that the R1S4W2 waveform increased prediction accuracy for test samples 1-4 by ~20% compared to the OG waveform. We found that the agnostic behavior towards pH was reproducible for R1S4W2 and not the OG. However, we did notice that the T3 a.c. challenge sample accuracy was not reproducible across electrodes for either waveform. We attribute this to variations in electrode fabrication. Standardizing the fabrication of fast voltammetry electrodes, along with multi-objective optimization with reproducibility as a metric, will help alleviate this issue. Regardless, the performance of R1S4W2 as an early optimization candidate showing enhanced test and challenge set accuracy demonstrates the success and future promise of the SeroOpt workflow.

### The SeroOpt workflow is reproducible and outperforms random search.

To investigate whether Bayesian optimization was "getting lucky" and not gleaning chemically relevant information, this process was repeated, starting with a new set of six random waveforms and carried out for four strings as described above

(**Fig. 4**). We refer to this as run 2 (R2). Note that data are aggregated across strings in each run; data between runs are not aggregated. The runs are kept separate to compare, from a new randomized initialization, whether four rounds of Bayesian optimization can repeatedly produce improved waveforms. We do not expect the convergence of R2 on the same waveform as R1. The search space is vast and given only the small subset of waveforms tested, converging on the same optima is unlikely. Instead, if R1 and R2 both find improved waveforms faster than the randomized waveforms, we can examine the black box models to see what makes the optimizer decide on 'good' waveforms (*vide infra*).

In all cases, except for the first run of pH and a.c. challenge samples, the average serotonin test/challenge set error was lower when using the optimized serotonin waveforms (W2,4,6,8 for S2,3,4 of R1 and R2), when compared to the average for the randomly generated S1 waveforms of R1 and R2 (**Fig. 4**). The error minima were lower in all cases for the optimized waveforms; random search never produced a better waveform than Bayesian optimization. For example, while each W2 in R1 improved across strings, R2S2W2 immediately found a 5-fold lower minimum than the starting initialization. Thus, new random initialization can cause the discovery of new waveforms in local minima.

These results suggest the following. On average, Bayesian optimization produces better waveforms than randomly generated or chemist-enabled waveforms. Bayesian optimization finds waveforms corresponding to error minima better than random chance. The Bayesian optimization surrogate model (*i.e.,* Gaussian process)

effectively models the relationship between voltammetry waveforms and performance, as the minima only occurred for waveforms optimized specifically for serotonin detection. In addition to outperforming random waveforms, waveforms optimized solely for dopamine did not outperform those optimized for serotonin when assessing serotonin output metric accuracy. For example, the average serotonin accuracy was ~45 nM using the randomly generated waveforms. By optimizing for any serotonin parameter (test set accuracy, a.c. accuracy, pH accuracy, detection limit), serotonin accuracy was improved to 34 nM (24% improvement). While an ostensibly small return on investment, this is only the first iteration of this protocol, and the results consistently outperformed the few standard alternatives to waveform design.

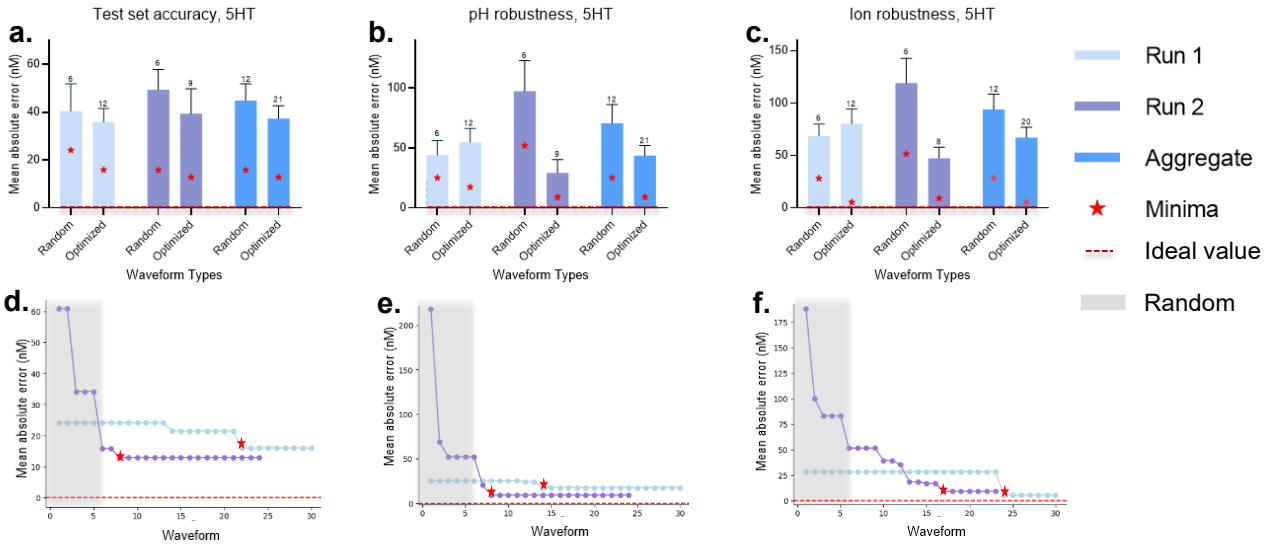



**Figure 4.** Bayesian optimization outperforms random search. Average mean absolute error for run 1, run 2 and the aggregate of both runs are shown for serotonin test set accuracy **(a)**, pH robustness **(b),** and ion robustness **(c)**. Error bars represent standard deviation. Sample size is shown atop the bars. The minima of error is for each group of waveforms is denoted by a red star. Random refers to string 1 waveforms only. Optimized refers to waveforms optimized for 5-HT performance (*i.e.,* W2,4,6,8). **(d-f)** Convergence plots corresponding to a-c, respectively, showing current minimum mean absolute error at each waveform iteration. Gray boxes represent random initialization waveform regions.

### Fine-grained waveform parameter tuning improves predictive performance.

In total, 55 waveforms were tested experimentally (the OG waveform, 12 randomly generated waveforms from R1S1 and R2S2, and 42 Bayesian optimized waveforms from R1 and R2 S2-4) with their corresponding metrics given as optimization training data. The generated waveforms covered a wide search space across all the waveform parameters. In **Figure 5**, clusters of points are interpreted as exploitation, while isolated points are interpreted as exploration. A key advantage of Bayesian optimization is that the acquisition function parsimoniously explores a search space with the exploration-exploitation trade-off in mind.[66] Bayesian

optimization judiciously explored the search space over 55 waveforms. At the time of writing and to our knowledge, this is the largest optimization scheme covered in neurochemical voltammetry waveform development.

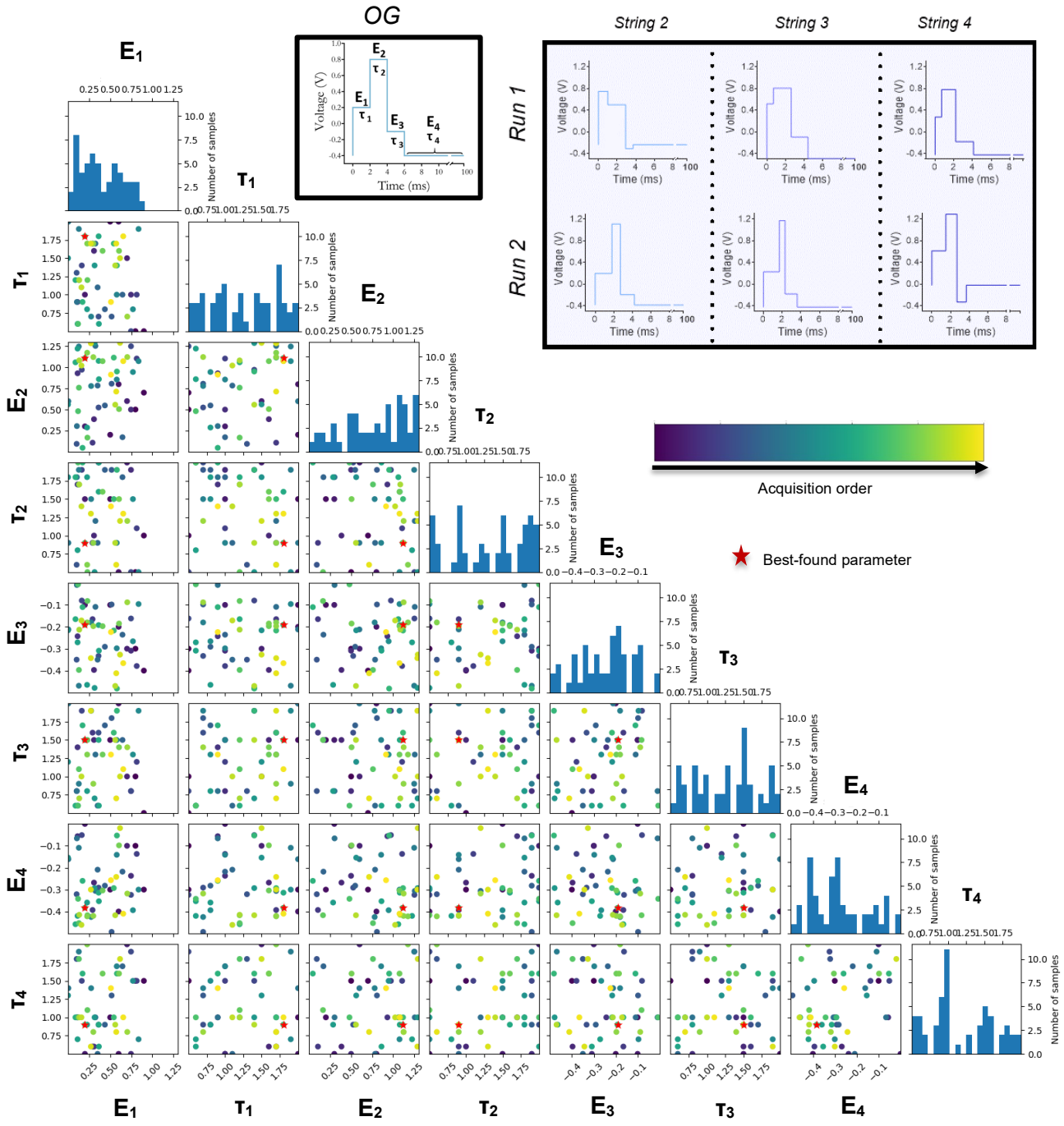



**Figure 5.** Search space of all waveforms tested experimentally from runs 1 and 2. Red star represents optimum parameters. Histograms represent the frequency of that parameter value in the waveforms tested. **(inset)** Evolution of the predicted Bayesian optimization waveforms across two separate Bayesian optimization runs, 1 and 2, for serotonin accuracy metric in blue (W2). String 1 not shown as they were randomly generated.

Data for all waveforms and metrics are provided (**Tables S2, S3**). We noticed that for serotonin accuracy (W2), the predicted waveforms between R1 and R2 looked similar, especially for S3 and S4 (**Fig. 5**, inset). The serotonin accuracy waveforms share characteristics with the OG waveform across R1 and R2. They exhibit low to high potential steps for the oxidative potential steps, and high to low potential steps for the reductive potential steps. By S4, all waveforms prefer the 'intermediate' anodic pulse step concept described in the VET literature, in which a relatively low amplitude $E_1$ step before a higher amplitude $E_2$ step prevents signal saturation and enhances concentration discrimination.[47] Further, most waveforms exhibited a large amplitude counter-pulse (*e.g.,* a large difference between $E_2$ and $E_3$ to complete the redox cycle).[63] The fact that the model is learning these domain knowledge heuristics across the four iterations suggests it can also learn more complex, higher-order interactions.

Waveform optimizations occurred with relatively small changes in E and τ, even for waveforms as simple as four steps, as shown here. Tuning these waveforms can result in dramatic improvements in the predictive performance differences of the resulting models. The effect of varying and reorganizing pulse parameters is relatively unexplored in a systematic, multi-variate manner, as done here. For example, R1S4W6 and R1S3W8 differed by ≤0.04 V and ≤0.9 ms in E and τ (**Table S2**). Yet, R1S3W8 outperformed R1S4W6 for serotonin test set, pH, and ion accuracy, with up to nearly a 50% reduction in error (**Table S3**).

To test whether this effect was due to differences in electrodes across strings (separate electrodes were used across strings to encourage generalizability across electrodes), we compared two similar waveforms tested on the same electrode: R2S1W2 and R2S1W3. These waveforms differed only by ≤0.15 V and ≤ 1.2 ms, yet R2S1W2 outperformed R2S1W3 in all serotonin metrics (**Table S2, S3**). Thus, small and otherwise "insignificant" changes in step potentials and holding times can produce significant accuracy differences. These findings support the importance of a technique like Bayesian optimization to tune parameters with fine-grained adjustments.

The order of the steps in the rapid pulse also matters. For example, R1S1W1 and R1S4W3 are nearly identical, except for the order of their pulses. Yet, R1S1W1 outperformed R1S4W3 in all categories up to five-fold (**Tables S2, S3**).

## Interpretable machine learning reveals waveform parameter interactions and learnable heuristics

Aside from the qualitative explanations above, interpretable machine learning methods[68] can be applied to 'open the black box' and assess how Bayesian optimization decides on improved waveforms. Thus, we investigated if the optimizer was learning the heuristics that electrochemists use to optimize waveforms, if it was learning novel relationships from the data, or both. We used a global, model-agnostic technique known as partial dependence plots (PDPs) to visualize how varying waveform parameters affect the surrogate model predictions.[68] The PDPs are useful

for non-parametric models, such as Gaussian processes, that are not directly interpretable.[68] Essentially, PDPs average the predictions from the model over samples where all parameters, except the ones of interest, are held constant. The effect of changing only the parameter(s) of interest can then be inferred (*i.e.,* the partial dependence of a feature).

The PDPs for the aggregated runs (R1 and R2 combined) and the individual runs are shown for the serotonin test set accuracy metric (**Figs. 6a, S5, S6**, respectively). We focus on the aggregated models because these have more total samples and, thus, are more likely to uncover meaningful relationships. The 2D plots on the diagonal represent the average effect of a metric while varying that parameter. Generally, the more a PDP plot for a particular feature varies, the more important that feature. Conversely, flat lines indicate either unimportant or interacting features.

The aggregated data PDPs (**Fig. 6a**) confirm a complex and interacting optimization landscape. For example, $E_3$ oscillates, $E_4$ is parabolic, and $E_1$ and $\tau_1$ are monotonically decreasing or increasing, respectively. The 3D contour plots below the diagonal represent the average effects on each metric while varying two waveform parameters. Because we minimize error, the purple shading represents the optimal (minima) regions, while the yellow regions represent maxima.

The PDPs have some weaknesses. First, PDPs represent averages, meaning heterogenous interactions can be obfuscated (*e.g.,* an effect on one-half of the data may be averaged out by an opposite effect on the other half). Thus, non-varying

parameters in PDPs could be misinterpreted. To confirm this, we examined individual conditional expectation (ICE) plots. The ICE plots show the individual contributions that make up the averages in the PDP plots.[68] Thus, the 2D PDPs (blue lines, **Fig. 6a**) have matching structure with the average ICE plots (blue lines, **Fig. 6b**). The individual instances (gray lines, **Fig. 6b**) show that there are heterogeneous effects hidden by the PDP averages for some parameters. For example, $\tau_1$, $E_3$, and $E_4$ have traces that do not all follow the same general trends. Thus, varying these parameters depends on heterogeneous interactions between the other waveform parameters. Meanwhile, the remaining parameters, $E_1$, $E_2$, $\tau_2$, $\tau_3$, and $\tau_4$, follow the same general trends (flat lines suggesting non-interacting waveform parameters).

As an alternative to PDP and ICE plots, we used Shapley additive explanations (SHAP) plots.[68] The SHAP values enable interpretations of how features contribute to individual model predictions. The SHAP plots confirmed that $E_3$, $E_4$, $\tau_1$, and $E_1$ were the most important features. **Figure 6c** shows the spread of the SHAP value per feature. Further, the heterogeneous effects, particularly in $E_3$ and $E_4$, are confirmed by the different colors of the feature values that do not cluster on a single side.

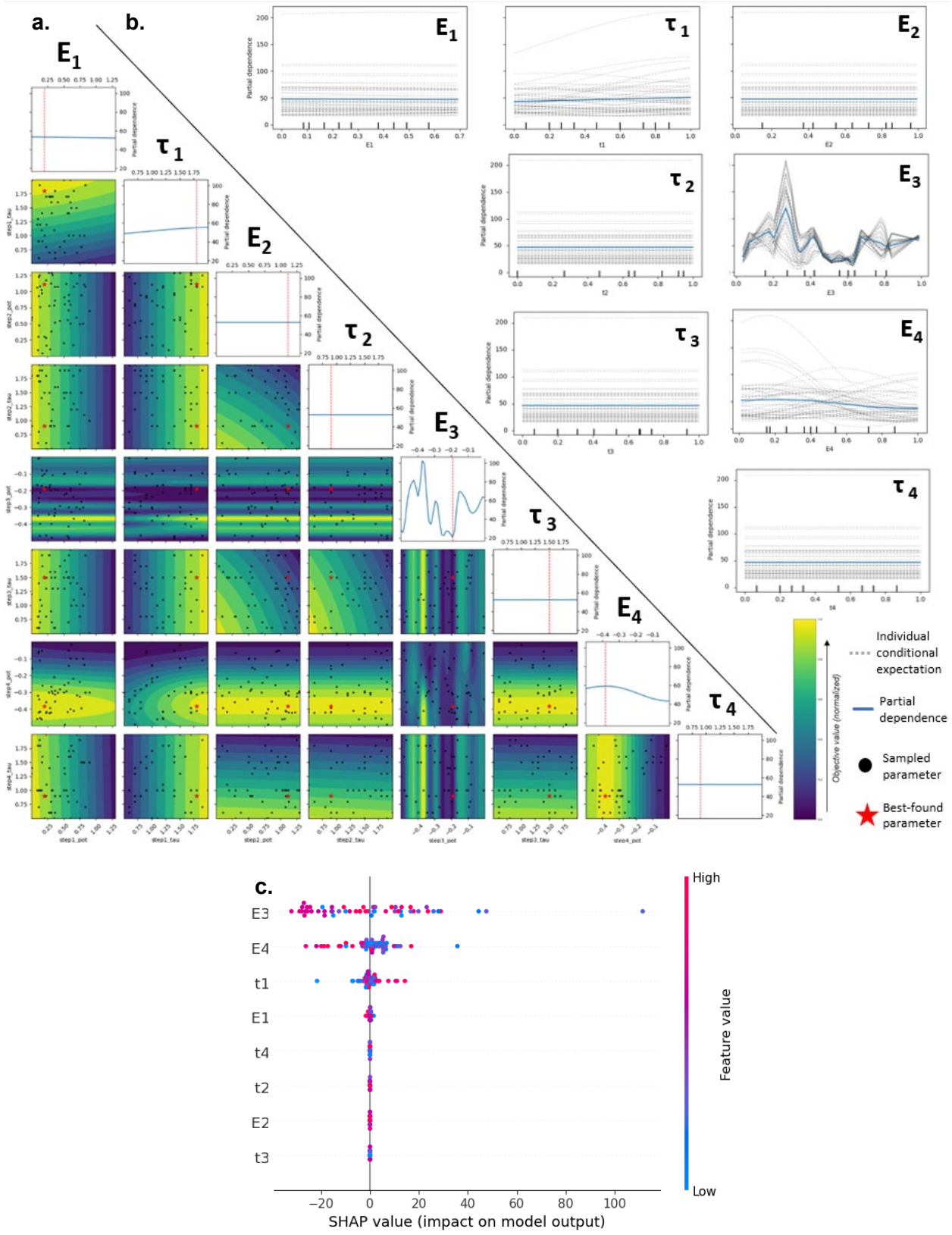



**Figure 6. a)** Partial dependence plot. **b)** Individual conditional expectation plots. Ticks represent deciles of the feature values. **c)** Shapley additive explanations summary plot.

# Discussion

Bayesian optimization enables global optima in high-dimensional search spaces to be identified by data-driven experimental designs across complex interaction parameters.[69] Bayesian optimization has been widely applied to diverse fields, including automated machine learning,[70] robotics,[71] sensor design,[72] materials discovery,[73,74] and chemical reaction optimization.[75,76] Despite its advantages and versatility, Bayesian optimization has rarely been applied to analytical chemistry, and specifically electrochemistry.[77-80]

Other approaches can be used to design waveforms (*e.g.*, first principles, chemometric screening, design of experiments). However, these approaches suffer from limited computational complexity, an exponential number of experiments required to optimize each parameter, resource intensity (labor, time, materials, *etc.*), and the inability to account for confounding waveform parameter interactions.[81] Our attempts to use feature selection to identify critical waveform step potentials and lengths were confounded by the magnitude of the current response and the pulse pattern (**Fig. S7**). The difficulty in electrochemical waveform design arises partly because each pulse (voltage and step length) affects the state of the interface between the solution and the working electrode. This interface evolves during and between each pulse. The effect of an individual pulse depends not only on its characteristics (E and $\tau$) but also on prior pulses.

We introduced an experimental design framework to embed voltammetry waveforms and their corresponding electroanalytical performance into a Bayesian

optimization workflow. Rather than optimizing for an electrochemical response, the accuracy of the supervised regression models was optimized directly by including model accuracy metrics as the objectives. We explored which model metrics were optimizable by simultaneously performing parallel single-objective optimization loops across eight metrics (**Fig. 2**). We found that serotonin test set accuracy optimization was sample-efficient, reproducible, and outperformed domain-guided and randomly designed waveforms across multiple metrics (**Fig. 3**).

We showed that in two separately initialized optimization campaigns consisting of four strings or 'rounds', we generated waveforms selective for serotonin in the presence of interferents (**Fig. 4**). Previous applications of Bayesian optimization in other fields achieved improvement in as few as three or four string-like iterations (*i.e.,* low data regimes). Thus, the behavior we observed was expected.[73,75,82,83] Notably, our selectivity challenges were more arduous and efficient than many standard waveform validation schemes that test only a single interferent or interferent concentration after a waveform is developed for an analyte of interest.

Selectivity is a significant barrier to effective waveform design, especially for background-inclusive and multi-analyte waveforms. Most voltammetry approaches achieve selectivity by either training a machine-learning model, modifying the waveform, or changing the electrode material. Rather than independently adopting one of the latter approaches, our data-driven waveform design uses the predictive performance of a machine learning model as feedback to modify waveform

parameters–the black box model decides what waveform would generate accurate PLSR predictions.

In addition to 5-HIAA, DOPAC, and ascorbate, monovalent cation concentrations (*i.e.*, $Na^+$, $K^+$, $H^+$) fluctuate in the brain extracellular space with neural stimulation due to the biophysics of membrane polarization and repolarization, transporter dynamics, and elevated $O_2$ consumption (and $CO_2$/carbonic acid/$H^+$ production) associated with synchronized action potentials.[84] Thus, these species represent key interferents to test in the presence of analytes, as electrodes will likely encounter changes in interferent concentrations under real-world (*in vivo*) conditions.

The literature suggests that specific voltage pulses can deconvolute monoamine neurotransmitter responses from cation changes.[85-87] Thus, we hypothesized our search space would contain interferent agnostic waveforms. We expected to find waveforms whose voltammograms, modeled in low-dimensional space by PLSR, are selective for features specific only to the analytes of interest (dopamine and serotonin) and not those affected by interferents. Training across such interferents is unnecessary if a waveform-model combination can ignore cation interferent effects (*i.e.,* is cation agnostic). Thus, we implicitly built the search for agnostic waveforms into our Bayesian workflow by introducing the concept of a challenge set.

Challenge set samples illustrated that SeroOpt can identify implicitly (*i.e.,* requiring no explicit training samples) interferent agnostic waveforms (**Fig. 3a**). While the literature has demonstrated cationic interferent agnostic waveforms,[67,85-

[87] our approach required no manual or additional data processing, and instead automatically acquired agnostic waveforms. Combining the information content of an optimized waveform with a powerful machine learning model (PLSR) enabled this agnostic behavior.

Because step potential,[47,63] step order,[46] and hold time[88] or hold potential[86] can impact waveform performance, other pulse techniques that layer steps at constant potentials and times could maximize their performance by tuning these parameters similarly to the manner presented here.[48] Adding more pulses could deteriorate model performance, as useless steps add noise to the data.[48] Thus, careful selection of the number of steps is paramount. We confirmed this by noting performance differences across waveforms with only slight parameter differences. We attribute this behavior to the unique faradaic and non-faradaic processes occurring at sub-ms timescales.[67,85,87,89]

Optimization of individual pulse step lengths results in different transient redox responses from the preceding pulses to become the starting state for the succeeding pulses, as opposed to letting the current decay to steady-state. A non-steady-state approach has been shown to discriminate compounds more efficiently using VETs. Yet, a lack of methods for optimizing individual step lengths has prevented wide adoption of this practice. Differentiating dopamine from norepinephrine has been accomplished using pulses with differences as small as 0.1 V, though without systematic design patterns.[90]

Potential mechanisms underlying interferent agnostic waveforms include diffusion layer depletion of the interfering species by the onset pulse ($E_1/\tau_1$),[91] and other differentiating information provided by unique pulse sequences and transient responses of the rapid pulses to the model.[85,87,88] More optimization campaigns, interpretability techniques, and the use of numerical simulation of species at electrode surfaces could uncover the phenomena at play.

Regardless, the finding that interferent agnostic waveforms can be identified and optimized, especially when forgoing background subtraction, shows the utility of historically categorized "nonspecific" capacitive currents. These findings show that analyte-specific information from appropriately designed waveforms occurs in the background current. This information is captured by our model without explicit training, even in the presence of interferents that affect the double layer. Previous reports have shown that pH and $Na^+/K^+$ fluxes can cause hundreds to thousands of nM prediction errors *in vitro*.[85,92] For the same fluxes, our waveform-model combinations show only tens of nM error or less, and do not require explicit training, specialized waveform augmentation, or data analysis.

We noticed that across runs and interpretability methods, $E_1$ or $\tau_1$ (onset pulse/time), $E_2$ and $E_3$ (pulse/counter pulse[63]), and $E_4$ (holding potential) were repeatedly ranked as the most important features for the surrogate models of serotonin test set accuracy. These parameters represent four known heuristics: counter pulse potential ($E_3$, useful for analyte confirmation),[63] holding potential ($E_4$, for analyte accumulation, sensitivity, and reduced serotonin fouling),[35] and $\tau_1$ and $E_1$

(onset time/intermediate potential; useful for selectivity and diffusion layer depletion).[91] Further meta-analyses of these behaviors will provide essential insights into unexpected electrochemical optimization design patterns.

Small amplitude onset pulses have been shown to improve the deconvolution and differentiation of ions such as $H^+$,[87] $Na^+$, and $K^+$,[85] along with small amplitude onset sweeps for drift and pH.[67,93] Again, carefully designed waveform tuning can result in explicit and implicit interferent agnostic waveforms. Other waveform parameters deemed unimportant in this study might be associated with the imposed constraints affecting the full exploration of parameter space or our relatively small sample size. Further, the interpretability methods are also estimates of the surrogate model, which itself is an estimate. Thus, our interpretations must be taken lightly as correlations, not causation.

The SeroOpt paradigm is immediately extendable to more than four steps (eight parameters) to create more complex waveforms. Future research into other optimization metrics, supervised regression and surrogate models/kernels, and additional analytes is underway.[94,95] For example, pulses have been shown to differentiate norepinephrine from dopamine.[90]

We note the extendibility of our waveform embedding approach. This embedding can be used for any waveform type, such as sweeps, where the parameter values represent the slope (scan rate) of each segment, along with parameters for start and stop potentials. Pulse and sweep designs can also be combined.[91] Similar approaches could also extend to embedding AC voltammetry parameters (*e.g.*,

amplitude, phase).[96] Thus, rather than starting from a historic performer and exploring new waveforms one factor at a time, entirely new waveforms can be discovered *de novo*. Our approach will help accelerate waveform development for new single- and multi-analyte panels in environments that hinder selectivity or other difficult-to-optimize metrics. Further exploration of waveforms with agnostic behavior and multi-objective optimization is underway for co-detection waveforms. To aid other investigators in this pursuit, we provide data, tutorial code notebooks, and videos at https://github.com/csmova/SeroOpt, as well as our corresponding open-source voltammetry acquisition and analysis software at https://github.com/csmova/SeroWare.

To our knowledge, we report the first application of active learning to electrochemical waveform design. Our study represents one of the largest-scale investigations of neurochemical detection waveforms. Using a data-driven approach, we generated a waveform for serotonin detection that outperformed our expert-designed waveform and randomly generated waveforms across various metrics. We demonstrated the ability to search for interferent-agnostic waveforms using *a priori* design of 'challenge' samples. We attributed the success of SeroOpt to the efficient fine-grained tuning of voltage and temporal waveform parameters by Bayesian optimization, each having complex interaction effects. Lastly, we interpreted our model with three separate techniques to confirm our model was learning a representation of the waveform optimization landscape that aligned with heuristics and domain knowledge.

# Methods

## Chemicals

Serotonin (5-HT) hydrochloride (#H9523), dopamine (DA) hydrochloride (#H8502), 5-hydroxyindoleacetic acid (5-HIAA) (#H8876), 3,4-dihydroxyphenylacetic acid (DOPAC) (#850217), and ascorbic acid (#A92902) were purchased from Sigma-Aldrich (St. Louis, MO). Artificial cerebrospinal fluid (aCSF) solutions were prepared as previously described.[43,97] The aCSF solution was adjusted on the day of each experiment to pH 7.1, 7.2, or 7.3 ± 0.03 using HCl (Fluka, #84415). Altered cation (a.c.) aCSF buffer contained the following ion composition: 31 mM NaCl (#73575), 120 mM KCl (#05257), 1.0 mM $NaH_2PO_4$ (#17844), 2.5 mM $NaHCO_3$ (#88208) purchased from Honeywell Fluka (Charlotte, NC), and 1.0 mM $CaCl_2$ (#499609) and 1.2 mM $MgCl_2$ (#449172) purchased from Sigma-Aldrich. All aqueous solutions were prepared using Milli-Q grade or higher water (Sigma-Aldrich).

## Electrode fabrication and polymerization

Carbon fiber microelectrodes were fabricated by vacuum-aspirating 7-µm diameter carbon fibers (T650/35, Cytec Carbon Fiber) into O.D. 1.2 mm x I.D. 0.69 mm, 10 cm length borosilicate glass capillaries (Sutter Instrument Company, Novato, CA, B120-69-10). A micropipette puller (P-1000, Sutter Instrument Company, Novato, CA) was used to pull each capillary into two electrodes tapering and sealing the glass around the carbon fiber. Four-part epoxy (Sigma Aldrich, Spurr Low Viscosity Embedding

Kit- EM0300) was backfilled into the tip of each electrode. Epoxied electrodes were dried at 70 °C for 8-12 h. Electrode tips were cut to ~100 µm using micro-scissors under an inverted microscope. For electrical conduction, the electrodes were backfilled with a non-toxic metal alloy of gallium-indium-tin, Galinstan (Alfa Aesar, 14634-18). Bare copper wire (0.0253-in. diameter, Archor B22) was polished using a 600-grit polishing disc and inserted into working electrode capillaries to serve as the electrical connection to the potentiostat. Epoxy (Loctite EA 1C) was then placed around the top of each electrode to secure the Cu wire in place. The epoxy was cured for 24 h at room temperature.

Electrode tips were cleaned with HPLC-grade isopropanol (Sigma Aldrich #34863) for 10 min. Electrodes were then overoxidized by applying a static 1.4 V potential for 20 min.[98] Low-density EDOT:Nafion solution was made by first preparing a 40 mM EDOT stock; 100 µL of this stock was added to 200 µL of Nafion and diluted with 20 mL of acetonitrile.[20] A triangle waveform (1.5 V to -0.8 V to 1.5 V) was applied using a CHI Instruments Electrochemical Analyzer 15x at 100 mV/s to generate a PEDOT:Nafion coating on each electrode.

## *In vitro* experiments

Reference electrodes were made by placing 0.025-inch silver wire (A-M Systems, 783500) into bleach (5-10% sodium hypochlorite, Clorox, Oakland CA) for 10 minutes. Each reference electrode was rinsed with distilled water before being used in experiments. A flow cell (NEC-FLOW-1, Pine Research Instrumentation Inc.) was used to make measurements with a VICI air-actuated injector (220-0302H; VICI

Valco Instruments, Houston, TX). An HPLC pump by Dionex (Sunnyvale, California) pumped aCSF through the flow cell at a constant flow rate of 1.0 mL/min (**Fig. 7**).

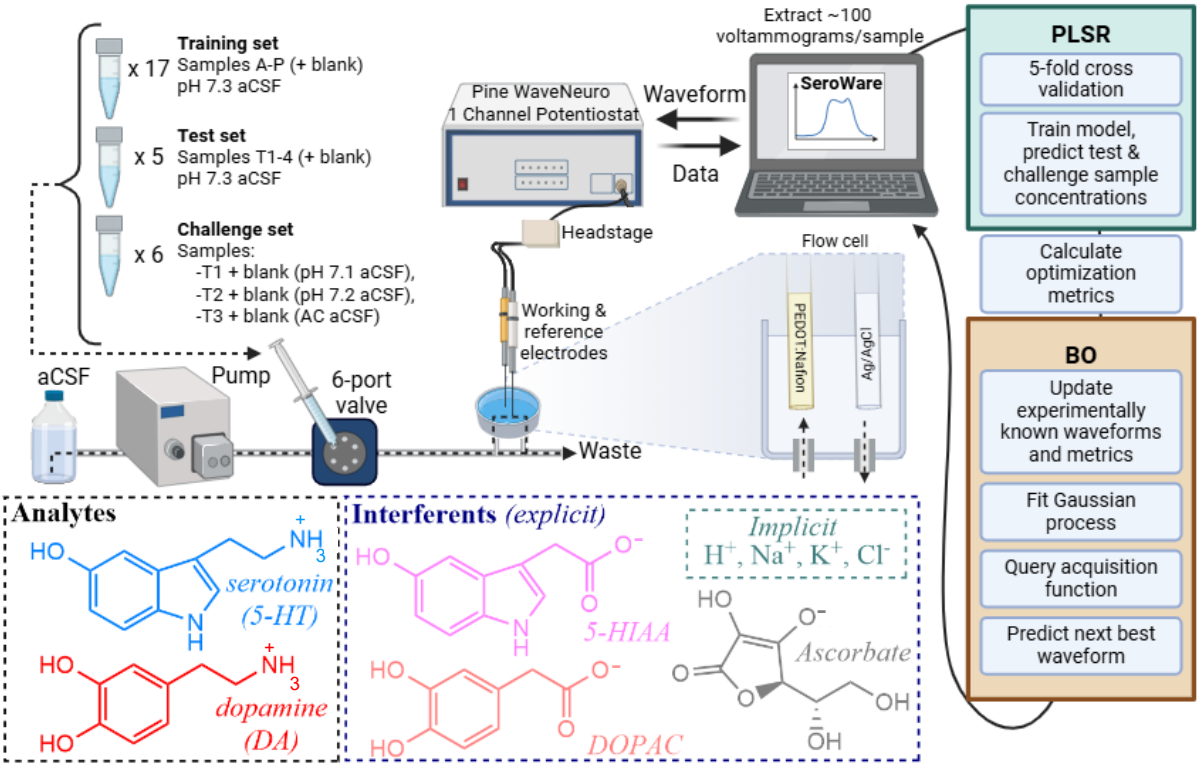



**Figure 7.** Workflow for parallel Bayesian optimization of voltammetric waveforms with intrinsic interferent selectivity.

Standard concentrations were selected using a fractional factorial box design (**Table 1**). This is a chemometric approach that designs a multi-dimensional 'box' spanning analytes, their concentrations, and experimental conditions of interest.[81,99] We selected a fractional approach to bias towards low analyte concentrations and small relative changes. High accuracy and precision in the nM range are important for monitoring basal and stimulated neurotransmitter levels using a single technique.

**Table 1.** Training, test and challenge set concentrations, in order of injection. All solutions prepared in artificial cerebrospinal fluid; a.c. = altered cations.

| Set | Sample | DA (nM) | 5-HT (nM) | 5-HIAA (µM) | DOPAC (µM) | Ascorbate (µM) | pH (units) | KCl (mM) | NaCl (mM) |
|---|---|---|---|---|---|---|---|---|---|
| Training | Blank | 0 | 0 | 0 | 0 | 0 | 7.3 | 3.5 | 147 |
| | A | 300 | 0 | 6 | 80 | 200 | 7.3 | 3.5 | 147 |
| | B | 1000 | 20 | 10 | 70 | 110 | 7.3 | 3.5 | 147 |
| | C | 0 | 120 | 6 | 90 | 190 | 7.3 | 3.5 | 147 |
| | D | 450 | 350 | 4 | 0 | 130 | 7.3 | 3.5 | 147 |
| | E | 600 | 500 | 1 | 10 | 170 | 7.3 | 3.5 | 147 |
| | Blank | 0 | 0 | 0 | 0 | 0 | 7.3 | 3.5 | 147 |
| | F | 160 | 250 | 2 | 20 | 180 | 7.3 | 3.5 | 147 |
| | G | 700 | 300 | 0 | 0 | 100 | 7.3 | 3.5 | 147 |
| | H | 80 | 160 | 10 | 60 | 100 | 7.3 | 3.5 | 147 |
| | I | 20 | 60 | 0 | 50 | 160 | 7.3 | 3.5 | 147 |
| | J | 40 | 40 | 2 | 100 | 120 | 7.3 | 3.5 | 147 |
| | Blank | 0 | 0 | 0 | 0 | 0 | 7.3 | 3.5 | 147 |
| | K | 800 | 10 | 8 | 30 | 150 | 7.3 | 3.5 | 147 |
| | L | 500 | 0 | 0 | 0 | 100 | 7.3 | 3.5 | 147 |
| | M | 0 | 250 | 0 | 0 | 100 | 7.3 | 3.5 | 147 |
| | N | 0 | 0 | 10 | 0 | 100 | 7.3 | 3.5 | 147 |
| | O | 0 | 0 | 0 | 50 | 100 | 7.3 | 3.5 | 147 |
| | P | 0 | 0 | 0 | 0 | 100 | 7.3 | 3.5 | 147 |
| | Blank | 0 | 0 | 0 | 0 | 0 | 7.3 | 3.5 | 147 |
| Test | T1 | 750 | 50 | 1 | 85 | 200 | 7.3 | 3.5 | 147 |
| | T2 | 100 | 400 | 5 | 9 | 200 | 7.3 | 3.5 | 147 |
| | T3 | 400 | 200 | 5 | 85 | 190 | 7.3 | 3.5 | 147 |
| | T4 | 70 | 30 | 5 | 35 | 200 | 7.3 | 3.5 | 147 |
| | Blank | 0 | 0 | 0 | 0 | 0 | 7.3 | 3.5 | 147 |
| Challenge (pH) | T1 pH | 750 | 50 | 1 | 85 | 200 | 7.1 | 3.5 | 147 |
| | Blank pH | 0 | 0 | 0 | 0 | 0 | 7.1 | 3.5 | 147 |
| | T2 pH | 100 | 400 | 5 | 9 | 200 | 7.2 | 3.5 | 147 |
| | Blank pH | 0 | 0 | 0 | 0 | 0 | 7.2 | 3.5 | 147 |
| Challenge (a.c.) | T3 a.c. | 400 | 200 | 5 | 85 | 190 | 7.3 | 120 | 31 |
| | Blank a.c. | 0 | 0 | 0 | 0 | 0 | 7.3 | 120 | 31 |

The fractional approach avoids a full factorial design, which would require orders of magnitude (and prohibitively) more calibration samples. In contrast, traditional calibration sets are information-poor and can lead to spurious correlations when

training a multiplexed method with overlapping signals from analytes and interferents.[81] The training and test sets effectively spanned the concentrations and combinations of analytes of interest without correlation (**Fig. S8**). Ascorbate was included in all samples (except blanks) for antioxidant properties. The concentrations of dopamine, serotonin, 5-HIAA, DOPAC, and ascorbate were altered over physiologically relevant changes in concentration throughout so the model could be trained and tested across all analytes.

Solutions of aCSF were purged with nitrogen for at least ten minutes before sample preparation. All training and test samples were prepared from stocks stored at -80 ˚C on the day of experiments. All solutions were adjusted to the corresponding pH each day prior to aliquoting. All solutions were kept covered from light and on ice during the experiments.

We define a training set (*i.e.*, calibration set) as known concentration analyte mixtures, *i.e.,* "standards", used to train a PLSR model. A test set is defined as known concentration analyte mixtures that were not used during training but instead held out and used to measure model performance. Test set samples only include samples with conditions occurring in the training set (*i.e.*, the same buffer conditions). We define "challenge" samples as additional test set samples prepared under conditions not included or varied in the training set, such as varied pH and cationic buffer salt concentrations (**Table 1**; see *Data Analysis*). We define an injection blank or zero (0) as an injected solution containing only aCSF.

Training, test, and challenge sets were injected (~1 mL into a 500-µL loop) into a flow cell using a six-port valve (**Fig. 7**). The valve was switched to the inject position for ~20 s per injection. The time between injections was ≥ 200 s, depending on the waveform and time for the current to return to baseline. Samples were injected in a pseudo-randomized but consistent order. Within each string, the waveform calibration curves were completed across consecutive days. All waveforms within a string were acquired with the same electrode. A different electrode was used for each string to ensure the robustness of the waveform optimization. All waveforms were conditioned for ≥10 min in aCSF before acquiring data.

## Voltammetry hardware and software

A two-electrode configuration *via* an Ag/AgCl reference electrode and a carbon fiber microelectrode working electrode was used. A PC with a PCIe-6363 data acquisition card (National Instruments (NI), Austin, TX) was used to control a WaveNeuro One FSCV Potentiostat System (NEC-WN-BASIC, Pine Research Instrumentation Inc.) with a 1,000 nA/V headstage amplifier (AC01HS2, Pine Research Instrumentation Inc.). The copper wire of the working electrode and the silver wire reference electrode were inserted into a microelectrode-headstage coupler (AC01HC0315-5, Pine Research Instrumentation Inc.) that connected the electrodes to the potentiostat.

In-house software was developed for RPV as described in a previous publication.[43] The software has since been updated and named SeroWare. Details regarding Seroware and open-source access to this software are found in a different

manuscript, titled 'SeroWare: An open source, end-to-end software suite for voltammetric acquisition and analysis of neurotransmitters' currently under review.

## Bayesian optimization

Bayesian optimization was carried out using the open-source Python package scikit-optimize.[100] This software uses an 'ask and tell' interface. First, the search space was constrained, as described in the *Results*. The surrogate model (Gaussian process regressor with a Matérn and white noise kernel, and uniform prior) was initialized through the 'tell' interface using vectorized and normalized String 1 waveform parameters and optimization metrics. A Matérn kernel was chosen because of its flexibility and the assumption that the true objective function of the waveform parameters is not infinitely differentiable (*i.e.*, the potentials and time applied by the potentiostat/data acquisition card are discretized to some degree).

The acquisition function (expected improvement) was then minimized using the 'ask' interface to generate a vectorized waveform to be experimentally queried. Kernel hyperparameters (*i.e.,* length scale, smoothness) and the acquisition function were optimized automatically by the limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm in the software package. The acquisition function returned a vectorized waveform that was then created in SeroWare format for data acquisition. After experimental results were obtained with the predicted waveform. The metrics of all previous waveforms were aggregated with the newest metrics. The

Bayesian optimizer was updated using the 'tell' interface to set new query points using the 'ask' interface.

In this work, increments of voltage were rounded to the nearest 0.001 V, and increments of time were rounded to the nearest 0.1 ms. Built-in partial dependence functions to scikit-learn and scikit-optimize were used to interpret the model, along with the SHAP Python package.

## Data analysis

Data were extracted using in-house custom acquisition software written in MATLAB 2016a. Models were built as described in previous literature using open-source Python packages (scikit-learn).[43,101] Briefly, roughly 40-100 voltammograms were extracted per sample injection. All voltammograms were normalized, and the number of components was chosen using 5-fold cross-validation. Optimization metrics were then calculated using the final model (**Table S1**).

***Drift training:*** The PLSR model was trained to account for drift using voltammograms collected throughout the experiment while aCSF containing interferents flowed and injections were not occurring (~2 h). We define these voltammograms as "background blanks". They are portions of the data when no samples are being injected. The injection blanks correct for injection artifacts, while the background blanks correct for drift (**Fig. S9**). Data in which drift was evident were extracted from these background epochs and labeled as 'zero' analyte concentrations to teach the model what drift, as opposed to analyte-containing, voltammograms,

looked like. Background blanks were in addition to data from injections of aCSF alone (*i.e.*, injection blanks), which accounted for flow cell injection artifacts.

We found this process increased the accuracy and precision of the PLSR predictions. It was generalizable to test set samples. We attribute this to a low-dimensional representation of drift learned by the model (**Fig. S9**). All concentration predictions were constrained to be ≥0 (*i.e.*, domain knowledge dictates concentrations cannot be negative). Negative concentration predictions were replaced with 0).

***Optimization metrics:*** The eight different optimization metrics were dopamine accuracy (mean absolute error of the test set predictions), serotonin accuracy (mean absolute error of the test set predictions), variance of the test set blanks (proxy for LOD) for zero dopamine or serotonin, mean absolute error for dopamine or serotonin in altered cation (a.c.) aCSF (ion robustness challenge samples), and varying pH aCSF (pH robustness challenge samples) (**Table S1**). Due to experimental time constraints, the LOD metric was excluded from the optimization panel for the second run of Bayesian optimization (R2). This resulted in 30 unique waveforms for the first run (six random waveforms in String 1, plus three strings of eight waveforms from subsequent rounds of Bayesian optimization) and a total of 24 waveforms for the second run (six random waveforms in String 1, plus six waveforms in three rounds of optimization). In R1 and R2 combined, 55 unique waveforms were tested (with the additional OG RPV waveform also tested; **Table S2**).

***Challenge samples:*** Test samples (T1-T4), prepared at pH 7.3, were used to assess dopamine and serotonin accuracy and LOD. Some test samples (T1-T3) were also prepared in aCSF at pH 7.1 or pH 7.2, and in aCSF with altered cation concentrations ($Na^+$ and $K^+$) to assess the accuracy of dopamine and serotonin predictions in the presence of varying $H^+$, $Na^+$, and $K^+$ concentrations expected *in vivo*. We refer to these specially prepared test samples as 'challenge' samples (**Table 1**, **Fig. 7**). These samples enabled sparse training set size. Thus, we could optimize for interferents agnostic waveforms without explicitly training on these interferents. Otherwise, training across variations in pH or other cations would require partial or up to full-fold increases in the samples injected. As an efficient alternative, we optimized for accuracy on the challenge set samples without any increase in training set size. Thus, the optimization goal of challenge samples was to find a waveform inherently agnostic to changes in pH or cations rather than a waveform that was 'trainable' across these interferents. In this case, the interferents implicitly optimized were pH and monovalent cations, which is extendable to any *a priori* domain knowledge of interferents expected. This approach is particularly useful in situations where the training data matrix differs from the model's application (*i.e., in vitro* to *in vivo* generalizability).

## Author Contributions

AMA, ASM, CL, CSM, KAP, and MAF conceived the work and designed the experiments. ANN, CSM, KKN, KAP, MEC, and MEW performed all experiments. CSM and KAP analyzed the data. CSM wrote the code for the regression and Bayesian optimization models. KAP performed statistical analyses. AMA, ASM, MAF, and CL guided the project. All authors wrote and approved the final version of the manuscript.

**Acknowledgments:** The authors acknowledge Biorender.com and Craiyon.com for figure generation.

**Funding:** This work was supported by the National Science Foundation (CHE-2404470). CSM was supported by the National Science Foundation Graduate Research Fellowship Program (DGE-1650604 and DGE-2034835). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

# References


1. Pang, Q.; Meng, J.; Gupta, S.; Hong, X.; Kwok, C. Y.; Zhao, J.; Jin, Y.; Xu, L.; Karahan, O.; Wang, Z.; Toll, S.; Mai, L.; Nazar, L. F.; Balasubramanian, M.; Narayanan, B.; Sadoway, D. R., Fast-charging aluminium–chalcogen batteries resistant to dendritic shorting. *Nature* **2022,** *608* (7924), 704-711. https://doi.org/10.1038/s41586-022-04983-9
2. Dinh Khac, H.; Whang, G.; Iadecola, A.; Makhlouf, H.; Barnabé, A.; Teurtrie, A.; Marinova, M.; Huvé, M.; Roch-Jeune, I.; Douard, C.; Brousse, T.; Dunn, B.; Roussel, P.; Lethien, C., Nanofeather ruthenium nitride electrodes for electrochemical capacitors. *Nat. Mater.* **2024**. https://doi.org/10.1038/s41563-024-01816-0
3. Garrido-Barros, P.; Derosa, J.; Chalkley, M. J.; Peters, J. C., Tandem electrocatalytic n2 fixation via proton-coupled electron transfer. *Nature* **2022,** *609* (7925), 71-76. https://doi.org/10.1038/s41586-022-05011-6
4. Patil, J. J.; Wan, C. T.-C.; Gong, S.; Chiang, Y.-M.; Brushett, F. R.; Grossman, J. C., Bayesian-optimization-assisted laser reduction of poly (acrylonitrile) for electrochemical applications. *ACS nano* **2023,** *17* (5), 4999-5013.
5. Zhang, W.; Lu, L.; Zhang, W.; Wang, Y.; Ware, S. D.; Mondragon, J.; Rein, J.; Strotman, N.; Lehnherr, D.; See, K. A.; Lin, S., Electrochemically driven cross-electrophile coupling of alkyl halides. *Nature* **2022,** *604* (7905), 292-297. https://doi.org/10.1038/s41586-022-04540-4
6. Greenhalgh, J. C.; Fahlberg, S. A.; Pfleger, B. F.; Romero, P. A., Machine learning-guided acyl-acp reductase engineering for improved in vivo fatty alcohol production. *Nat. Commun.* **2021,** *12* (1), 5825. 10.1038/s41467-021-25831-w
7. Romero, P. A.; Krause, A.; Arnold, F. H., Navigating the protein fitness landscape with Gaussian processes. *Proc. Natl. Acad. Sci. U.S.A.* **2013,** *110* (3), E193-E201. doi:10.1073/pnas.1215251110
8. Li, J.; Liu, Y.; Yuan, L.; Zhang, B.; Bishop, E. S.; Wang, K.; Tang, J.; Zheng, Y.-Q.; Xu, W.; Niu, S.; Beker, L.; Li, T. L.; Chen, G.; Diyaolu, M.; Thomas, A.-L.; Mottini, V.; Tok, J. B. H.; Dunn, J. C. Y.; Cui, B.; Pașca, S. P.; Cui, Y.; Habtezion, A.; Chen, X.; Bao, Z., A tissue-like neurotransmitter sensor for the brain and gut. *Nature* **2022,** *606* (7912), 94-101. https://doi.org/10.1038/s41586-022-04615-2
9. Willmore, L.; Cameron, C.; Yang, J.; Witten, I. B.; Falkner, A. L., Behavioural and dopaminergic signatures of resilience. *Nature* **2022,** *611* (7934), 124-132. https://doi.org/10.1038/s41586-022-05328-2
10. Batten, S. R.; Bang, D.; Kopell, B. H.; Davis, A. N.; Heflin, M.; Fu, Q.; Perl, O.; Ziafat, K.; Hashemi, A.; Saez, I.; Barbosa, L. S.; Twomey, T.; Lohrenz, T.; White, J. P.; Dayan, P.; Charney, A. W.; Figee, M.; Mayberg, H. S.; Kishida, K. T.; Gu, X.; Montague, P. R., Dopamine and serotonin in human substantia nigra track social context and value signals during economic exchange. *Nat. Hum. Behav.* **2024**. https://doi.org/10.1038/s41562-024-01831-w
11. Flagel, S. B.; Clark, J. J.; Robinson, T. E.; Mayo, L.; Czuj, A.; Willuhn, I.; Akers, C. A.; Clinton, S. M.; Phillips, P. E. M.; Akil, H., A selective role for dopamine in stimulus–reward learning. *Nature* **2011,** *469* (7328), 53-57. https://doi.org/10.1038/nature09588
12. Phillips, P. E. M.; Stuber, G. D.; Heien, M. L. A. V.; Wightman, R. M.; Carelli, R. M., Subsecond dopamine release promotes cocaine seeking. *Nature* **2003,** *422* (6932), 614-618. https://doi.org/10.1038/nature01476
13. Pipita, M. E.; Santonico, M.; Pennazza, G.; Zompanti, A.; Fazzina, S.; Cavalieri, D.; Bruno, F.; Angeletti, S.; Pedone, C.; Incalzi, R. A., Integration of voltammetric analysis, protein electrophoresis and pH measurement for diagnosis of pleural effusions: A non-conventional diagnostic approach. *Sci. Rep.* **2020,** *10* (1), 15222. https://doi.org/10.1038/s41598-020-71542-5

14. Zhao, S.; Li, H.; Dai, J.; Jiang, Y.; Zhan, G.; Liao, M.; Sun, H.; Shi, Y.; Ling, C.; Yao, Y.; Zhang, L., Selective electrosynthesis of chlorine disinfectants from seawater. *Nat. Sustain.* **2024,** *7* (2), 148-157. https://doi.org/10.1038/s41893-023-01265-8
15. Mutz, Y. S.; do Rosario, D.; Silva, L. R. G.; Galvan, D.; Janegitz, B. C.; de Q. Ferreira, R.; Conte-Junior, C. A., A single screen-printed electrode in tandem with chemometric tools for the forensic differentiation of Brazilian beers. *Sci. Rep.* **2022,** *12* (1), 5630. https://doi.org/10.1038/s41598-022-09632-9
16. Andrews, A. M., The BRAIN Initiative: Toward a chemical connectome. *ACS Chem. Neurosci.* **2013,** *4* (5), 645. https://doi.org/10.1021/cn4001044
17. Puthongkham, P.; Venton, B. J., Recent advances in fast-scan cyclic voltammetry. *Analyst* **2020,** *145* (4), 1087-1102. https://doi.org/10.1039/c9an01925a
18. Rodeberg, N. T.; Sandberg, S. G.; Johnson, J. A.; Phillips, P. E. M.; Wightman, R. M., Hitchhiker's guide to voltammetry: Acute and chronic electrodes for in vivo fast-scan cyclic voltammetry. *ACS Chem. Neurosci.* **2017,** *8* (2), 221-234. https://doi.org/10.1021/acschemneuro.6b00393
19. Daws, L.; Andrews, A.; Gerhardt, G., Electrochemical techniques and advances in psychopharmacology. In *Encyclopedia of psychopharmacology*, IP, S.; LH, P., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; pp 1-6.
20. Vreeland, R. F.; Atcherley, C. W.; Russell, W. S.; Xie, J. Y.; Lu, D.; Laude, N. D.; Porreca, F.; Heien, M. L., Biocompatible PEDOT:Nafion composite electrode coatings for selective detection of neurotransmitters in vivo. *Anal. Chem.* **2015,** *87* (5), 2600-2607. https://doi.org/10.1021/ac502165f
21. Shao, Z.; Chang, Y.; Venton, B. J., Carbon microelectrodes with customized shapes for neurotransmitter detection: A review. *Anal. Chim. Acta* **2022,** *1223*, 340165. https://doi.org/10.1016/j.aca.2022.340165
22. Castagnola, E.; Robbins, E. M.; Krahe, D. D.; Wu, B.; Pwint, M. Y.; Cao, Q.; Cui, X. T., Stable in-vivo electrochemical sensing of tonic serotonin levels using PEDOT/CNT-coated glassy carbon flexible microelectrode arrays. *Biosens. Bioelectron.* **2023,** *230*, 115242. https://doi.org/10.1016/j.bios.2023.115242
23. Rafi, H.; Zestos, A. G., Multiplexing neurochemical detection with carbon fiber multielectrode arrays using fast-scan cyclic voltammetry. *Anal. Bioanal. Chem.* **2021,** *413* (27), 6715-6726.
24. Swamy, B. E. K.; Venton, B. J., Carbon nanotube-modified microelectrodes for simultaneous detection of dopamine and serotoninin vivo. *Analyst* **2007,** *132* (9), 876-884. https://doi.org/10.1039/B705552H
25. Mena, S.; Visentin, M.; Witt, C. E.; Honan, L. E.; Robins, N.; Hashemi, P., Novel, user-friendly experimental and analysis strategies for fast voltammetry: Next generation FSCAV with artificial neural networks. *ACS Meas. Sci. Au* **2022,** *2* (3), 241-250. https://doi.org/10.1021/acsmeasuresciau.1c00060
26. Twomey, T.; Barbosa, L.; Lohrenz, T.; Montague, P. R., Deep learning architectures for FSCV, a comparison. *arXiv (Medical Physics)* **2022,** (posted 2022-12-05). https://doi.org/10.48550/arXiv.2212.01960 (accessed 2023-12-12)
27. Montague, P. R.; Kishida, K. T., Computational underpinnings of neuromodulation in humans. *Cold Spring Harbor Symp. Quant. Biol.* **2018,** *83*, 71-82. https://doi.org/10.1101/sqb.2018.83.038166
28. Dunham, K. E.; Venton, B. J., Improving serotonin fast-scan cyclic voltammetry detection: New waveforms to reduce electrode fouling. *Analyst* **2020,** *145* (22), 7437-7446. https://doi.org/10.1039/D0AN01406K
29. Rafi, H.; Zestos, A. G., Recent advances in FSCV detection of neurochemicals via waveform and carbon microelectrode modification. *J. Electrochem. Soc.* **2021,** *168* (5), 057520.

30. Fedorowski, J.; LaCourse, W. R., A review of pulsed electrochemical detection following liquid chromatography and capillary electrophoresis. *Anal. Chim. Acta* **2015,** *861*, 1-11. https://doi.org/10.1016/j.aca.2014.08.035
31. Wei, Z.; Yang, Y.; Wang, J.; Zhang, W.; Ren, Q., The measurement principles, working parameters and configurations of voltammetric electronic tongues and its applications for foodstuff analysis. *J. Food Eng.* **2018,** *217*, 75-92. https://doi.org/10.1016/j.jfoodeng.2017.08.005
32. Moro, G.; Silvestri, A.; Ulrici, A.; Conzuelo, F.; Zanardi, C., How to optimize the analytical performance of differential pulse voltammetry: One variable at time versus design of experiments. *J Solid State Electrochem.* **2024,** *28* (3), 1403-1415. 10.1007/s10008-023-05753-x
33. Jaworski, A.; Rapecki, T.; Wikiel, K., Consolidated designer waveform for maximizing analytical output of voltammetric measurements for complex chemical matrices. *J. Electroanal. Chem.* **2023,** *936*, 117332. https://doi.org/10.1016/j.jelechem.2023.117332
34. Hashemi, P.; Dankoski, E. C.; Petrovic, J.; Keithley, R. B.; Wightman, R. M., Voltammetric detection of 5-hydroxytryptamine release in the rat brain. *Anal. Chem.* **2009,** *81* (22), 9462-71. https://doi.org/10.1021/ac9018846
35. Venton, B. J.; Cao, Q., Fundamentals of fast-scan cyclic voltammetry for dopamine detection. *Analyst* **2020,** *145* (4), 1158-1168. https://doi.org/10.1039/C9AN01586H
36. Heien, M. L. A. V.; Phillips, P. E. M.; Stuber, G. D.; Seipel, A. T.; Wightman, R. M., Overoxidation of carbon-fiber microelectrodes enhances dopamine adsorption and increases sensitivity. *Analyst* **2003,** *128* (12), 1413-1419. https://doi.org/10.1039/B307024G
37. Kim, S. Y.; Oh, Y. B.; Shin, H. J.; Kim, D. H.; Kim, I. Y.; Bennet, K.; Lee, K. H.; Jang, D. P., 5-hydroxytryptamine measurement using paired pulse voltammetry. *Biomed. Eng. Lett* **2013,** *3* (2), 102-108. https://doi.org/10.1007/s13534-013-0093-z
38. Park, C.; Oh, Y.; Shin, H.; Kim, J.; Kang, Y.; Sim, J.; Cho, H. U.; Lee, H. K.; Jung, S. J.; Blaha, C. D.; Bennet, K. E.; Heien, M. L.; Lee, K. H.; Kim, I. Y.; Jang, D. P., Fast cyclic square-wave voltammetry to enhance neurotransmitter selectivity and sensitivity. *Anal. Chem.* **2018,** *90* (22), 13348-13355. https://doi.org/10.1021/acs.analchem.8b02920
39. Shin, H.; Oh, Y.; Park, C.; Kang, Y.; Cho, H. U.; Blaha, C. D.; Bennet, K. E.; Heien, M. L.; Kim, I. Y.; Lee, K. H.; Jang, D. P., Sensitive and selective measurement of serotonin in vivo using fast cyclic square-wave voltammetry. *Anal. Chem.* **2020,** *92* (1), 774-781. https://doi.org/10.1021/acs.analchem.9b03164
40. Oh, Y.; Heien, M. L.; Park, C.; Kang, Y. M.; Kim, J.; Boschen, S. L.; Shin, H.; Cho, H. U.; Blaha, C. D.; Bennet, K. E.; Lee, H. K.; Jung, S. J.; Kim, I. Y.; Lee, K. H.; Jang, D. P., Tracking tonic dopamine levels in vivo using multiple cyclic square wave voltammetry. *Biosens. Bioelectron.* **2018,** *121*, 174-182. https://doi.org/10.1016/j.bios.2018.08.034
41. Shin, H.; Goyal, A.; Barnett, J. H.; Rusheen, A. E.; Yuen, J.; Jha, R.; Hwang, S. M.; Kang, Y.; Park, C.; Cho, H.-U.; Blaha, C. D.; Bennet, K. E.; Oh, Y.; Heien, M. L.; Jang, D. P.; Lee, K. H., Tonic serotonin measurements in vivo using N-shaped multiple cyclic square wave voltammetry. *Anal. Chem.* **2021,** *93* (51), 16987-16994. https://doi.org/10.1021/acs.analchem.1c02131
42. Abdalla, A.; Atcherley, C. W.; Pathirathna, P.; Samaranayake, S.; Qiang, B.; Peña, E.; Morgan, S. L.; Heien, M. L.; Hashemi, P., In vivo ambient serotonin measurements at carbon-fiber microelectrodes. *Anal. Chem.* **2017,** *89* (18), 9703-9711. https://doi.org/10.1021/acs.analchem.7b01257
43. Movassaghi, C. S.; Perrotta, K. A.; Yang, H.; Iyer, R.; Cheng, X.; Dagher, M.; Fillol, M. A.; Andrews, A. M., Simultaneous serotonin and dopamine monitoring across timescales by rapid pulse voltammetry with partial least squares regression. *Anal. Bioanal. Chem.* **2021,** *413* (27), 6747-6767. https://doi.org/10.1007/s00216-021-03665-1

44. Movassaghi, C. S.; Alcañiz Fillol, M.; Kishida, K. T.; McCarty, G.; Sombers, L. A.; Wassum, K. M.; Andrews, A. M., Maximizing electrochemical information: A perspective on background-inclusive fast voltammetry. *Anal. Chem.* **2024**. https://doi.org/10.1021/acs.analchem.3c04938
45. Bard, A. J.; Faulkner, L. R.; White, H. S., *Electrochemical methods: Fundamentals and applications*. John Wiley & Sons: 2022.
46. Campos, I.; Alcañiz, M.; Masot, R.; Soto, J.; Martínez-Máñez, R.; Vivancos, J.-L.; Gil, L., A method of pulse array design for voltammetric electronic tongues. *Sens. Actuators, B* **2012,** *161* (1), 556-563. https://doi.org/10.1016/j.snb.2011.10.075
47. Fuentes, E.; Alcañiz, M.; Contat, L.; Baldeón, E. O.; Barat, J. M.; Grau, R., Influence of potential pulses amplitude sequence in a voltammetric electronic tongue (VET) applied to assess antioxidant capacity in aliso. *Food Chem.* **2017,** *224*, 233-241. https://doi.org/10.1016/j.foodchem.2016.12.076
48. Ivarsson, P.; Holmin, S.; Höjer, N.-E.; Krantz-Rülcker, C.; Winquist, F., Discrimination of tea by means of a voltammetric electronic tongue and different applied waveforms. *Sens. Actuators, B* **2001,** *76* (1), 449-454. https://doi.org/10.1016/S0925-4005(01)00583-4
49. Ross, A. E.; Venton, B. J., Sawhorse waveform voltammetry for selective detection of adenosine, ATP, and hydrogen peroxide. *Anal. Chem.* **2014,** *86* (15), 7486-7493. https://doi.org/10.1021/ac501229c
50. Altieri, S. C.; Yang, H.; O'Brien, H. J.; Redwine, H. M.; Senturk, D.; Hensler, J. G.; Andrews, A. M., Perinatal vs genetic programming of serotonin states associated with anxiety. *Neuropsychopharmacology* **2015,** *40* (6), 1456-70. https://doi.org/10.1038/npp.2014.331
51. Yang, H.; Thompson, A. B.; McIntosh, B. J.; Altieri, S. C.; Andrews, A. M., Physiologically relevant changes in serotonin resolved by fast microdialysis. *ACS Chem. Neurosci.* **2013,** *4* (5), 790-8. https://doi.org/10.1021/cn400072f
52. Altieri, S.; Singh, Y.; Sibille, E.; Andrews, A. M., Serotonergic pathways in depression. In *Neurobiology of Depression*, CRC Press: 2011; Vol. 20115633, pp 143-170.
53. Marcinkiewcz, C. A.; Mazzone, C. M.; D'Agostino, G.; Halladay, L. R.; Hardaway, J. A.; DiBerto, J. F.; Navarro, M.; Burnham, N.; Cristiano, C.; Dorrier, C. E.; Tipton, G. J.; Ramakrishnan, C.; Kozicz, T.; Deisseroth, K.; Thiele, T. E.; McElligott, Z. A.; Holmes, A.; Heisler, L. K.; Kash, T. L., Serotonin engages an anxiety and fear-promoting circuit in the extended amygdala. *Nature* **2016,** *537* (7618), 97-101. https://doi.org/10.1038/nature19318
54. Tye, K. M.; Prakash, R.; Kim, S.-Y.; Fenno, L. E.; Grosenick, L.; Zarabi, H.; Thompson, K. R.; Gradinaru, V.; Ramakrishnan, C.; Deisseroth, K., Amygdala circuitry mediating reversible and bidirectional control of anxiety. *Nature* **2011,** *471* (7338), 358-362. https://doi.org/10.1038/nature09820
55. Movassaghi, C. S.; Andrews, A. M., Call me serotonin. *Nat. Chem.* **2024,** *16* (4), 670-670. https://doi.org/10.1038/s41557-024-01488-y
56. Gershon, M. D.; Margolis, K. G., The gut, its microbiome, and the brain: Connections and communications. *The Journal of Clinical Investigation* **2021,** *131* (18). https://doi.org/10.1172/JCI143768
57. Murphy, D. L.; Fox, M. A.; Timpano, K. R.; Moya, P. R.; Ren-Patterson, R.; Andrews, A. M.; Holmes, A.; Lesch, K.-P.; Wendland, J. R., How the serotonin story is being rewritten by new gene-based discoveries principally related to SLC6A4, the serotonin transporter gene, which functions to influence all cellular serotonin systems. *Neuropharmacology* **2008,** *55* (6), 932-960. https://doi.org/10.1016/j.neuropharm.2008.08.034
58. Singh, Y. S.; Altieri, S. C.; Gilman, T. L.; Michael, H. M.; Tomlinson, I. D.; Rosenthal, S. J.; Swain, G. M.; Murphey-Corb, M. A.; Ferrell, R. E.; Andrews, A. M., Differential serotonin transport is linked to the rh5-HTTLPR in peripheral blood cells. *Translational Psychiatry* **2012,** *2* (2), e77-e77. https://doi.org/10.1038/tp.2012.2

59. Wrona, M. Z.; Dryhurst, G., Electrochemical oxidation of 5-hydroxytryptamine in aqueous solution at physiological pH. *Bioorg. Chem.* **1990,** *18* (3), 291-317. https://doi.org/10.1016/0045-2068(90)90005-P
60. Eltahir, A.; White, J.; Lohrenz, T.; Montague, P. R., Low amplitude burst detection of catecholamines. *bioRxiv (Neuroscience)* **2021,** (posted 2021-08-04). https://doi.org/10.1101/2021.08.02.454747 (accessed 2023-12-12)
61. Montague, P. R.; Lohrenz, T.; White, J.; Moran, R. J.; Kishida, K. T., Random burst sensing of neurotransmitters. *bioRxiv (Neuroscience)* **2019,** (posted 2019-04-13), 607077. https://doi.org/10.1101/607077 (accessed 2023-12-12)
62. Slautin, B. N.; Liu, Y.; Funakubo, H.; Vasudevan, R. K.; Ziatdinov, M.; Kalinin, S. V., Bayesian conavigation: Dynamic designing of the material digital twins via active learning. *ACS Nano* **2024,** *18* (36), 24898-24908. 10.1021/acsnano.4c05368
63. Campos, I.; Masot, R.; Alcañiz, M.; Gil, L.; Soto, J.; Vivancos, J. L.; García-Breijo, E.; Labrador, R. H.; Barat, J. M.; Martínez-Mañez, R., Accurate concentration determination of anions nitrate, nitrite and chloride in minced meat using a voltammetric electronic tongue. *Sens. Actuators, B* **2010,** *149* (1), 71-78. https://doi.org/10.1016/j.snb.2010.06.028
64. Roijers, D. M.; Vamplew, P.; Whiteson, S.; Dazeley, R., A survey of multi-objective sequential decision-making. *J. Artif. Intell. Res.* **2013,** *48*, 67-113.
65. Dagher, M.; Perrotta, K. A.; Erwin, S. A.; Hachisuka, A.; Ayer, R.; Masmanidis, S.; Yang, H.; Andrews, A. M., Optogenetic stimulation of midbrain dopamine neurons produces striatal serotonin release. *ACS Chem. Neurosci.* **2022,** *13* (7), 946-958.
66. Frazier, P. I., A tutorial on Bayesian optimization. *arXiv* **2018**. https://doi.org/10.48550/arXiv.1807.02811
67. Meunier, C. J.; Mitchell, E. C.; Roberts, J. G.; Toups, J. V.; McCarty, G. S.; Sombers, L. A., Electrochemical selectivity achieved using a double voltammetric waveform and partial least squares regression: Differentiating endogenous hydrogen peroxide fluctuations from shifts in pH. *Anal. Chem.* **2018,** *90* (3), 1767-1776. https://doi.org/10.1021/acs.analchem.7b03717
68. Molnar, C., *Interpretable machine learning*. Leanpub: https://christophm.github.io/interpretable-ml-book/, 2020.
69. Brochu, E.; Cora, V. M.; Freitas, N. d., A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *ArXiv* **2010,** *abs/1012.2599*. https://doi.org/10.48550/arXiv.1012.2599
70. Klein, A.; Falkner, S.; Bartels, S.; Hennig, P.; Hutter, F., Fast Bayesian optimization of machine learning hyperparameters on large datasets. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, Aarti, S.; Jerry, Z., Eds. PMLR: Proceedings of Machine Learning Research, 2017; Vol. 54, pp 528--536.
71. Calandra, R.; Seyfarth, A.; Peters, J.; Deisenroth, M. P., Bayesian optimization for learning gaits under uncertainty. *Ann. Math. Artif. Intell.* **2016,** *76* (1), 5-23. 10.1007/s10472-015-9463-9
72. Garnett, R.; Osborne, M. A.; Roberts, S. J., Bayesian optimization for sensor set selection. In *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks*, Association for Computing Machinery: Stockholm, Sweden, 2010; pp 209–219.
73. Wahl, C. B.; Aykol, M.; Swisher, J. H.; Montoya, J. H.; Suram, S. K.; Mirkin, C. A., Machine learning-accelerated design and synthesis of polyelemental heterostructures. *Sci. Adv.* **2021,** *7* (52), eabj5505. https://doi.org/10.1126/sciadv.abj5505
74. Liang, Q.; Gongora, A. E.; Ren, Z.; Tiihonen, A.; Liu, Z.; Sun, S.; Deneault, J. R.; Bash, D.; Mekki-Berrada, F.; Khan, S. A.; Hippalgaonkar, K.; Maruyama, B.; Brown, K. A.; Fisher Iii, J.; Buonassisi, T., Benchmarking the performance of Bayesian optimization across multiple

experimental materials science domains. *Npj Comput. Mater.* **2021,** *7* (1), 188. https://doi.org/10.1038/s41524-021-00656-9
75. Shields, B. J.; Stevens, J.; Li, J.; Parasram, M.; Damani, F.; Alvarado, J. I. M.; Janey, J. M.; Adams, R. P.; Doyle, A. G., Bayesian reaction optimization as a tool for chemical synthesis. *Nature* **2021,** *590* (7844), 89-96. https://doi.org/10.1038/s41586-021-03213-y
76. Griffiths, R.-R.; Hernández-Lobato, J. M., Constrained Bayesian optimization for automatic chemical design using variational autoencoders. *Chem. Sci.* **2020,** *11* (2), 577-586. 10.1039/C9SC04026A
77. Gundry, L.; Guo, S.-X.; Kennedy, G.; Keith, J.; Robinson, M.; Gavaghan, D.; Bond, A. M.; Zhang, J., Recent advances and future perspectives for automated parameterisation, Bayesian inference and machine learning in voltammetry. *Chem. Comm.* **2021,** *57* (15), 1855-1870. https://doi.org/10.1039/D0CC07549C
78. Bond, A. M., A perceived paucity of quantitative studies in the modern era of voltammetry: Prospects for parameterisation of complex reactions in Bayesian and machine learning frameworks. *J Solid State Electrochem.* **2020,** *24* (9), 2041-2050. https://doi.org/10.1007/s10008-020-04639-6
79. Puthongkham, P.; Wirojsaengthong, S.; Suea-Ngam, A., Machine learning and chemometrics for electrochemical sensors: Moving forward to the future of analytical chemistry. *Analyst* **2021,** *146* (21), 6351-6364. https://doi.org/10.1039/D1AN01148K
80. Fenton Jr, A. M.; Brushett, F. R., Using voltammetry augmented with physics-based modeling and Bayesian hypothesis testing to identify analytes in electrolyte solutions. *J. Electroanal. Chem.* **2022,** *904*, 115751. https://doi.org/10.1016/j.jelechem.2021.115751
81. Díaz-Cruz, J. M.; Esteban, M.; Ariño, C., *Chemometrics in electroanalysis*. 1 ed.; Springer: Cham, Switzerland, 2019; p 202.
82. Torres, J. A. G.; Lau, S. H.; Anchuri, P.; Stevens, J. M.; Tabora, J. E.; Li, J.; Borovika, A.; Adams, R. P.; Doyle, A. G., A multi-objective active learning platform and web app for reaction optimization. *J. Am. Chem. Soc.* **2022**. https://doi.org/10.1021/jacs.2c08592
83. Pruksawan, S.; Lambard, G.; Samitsu, S.; Sodeyama, K.; Naito, M., Prediction and optimization of epoxy adhesive strength from a small dataset through active learning. *Sci. Technol. Adv. Mate.* **2019,** *20* (1), 1010-1021. https://doi.org/10.1080/14686996.2019.1673670
84. Chesler, M.; Kaila, K., Modulation of pH by neuronal activity. *Trends Neurosci.* **1992,** *15* (10), 396-402. https://doi.org/10.1016/0166-2236(92)90191-A
85. Johnson, J. A.; Hobbs, C. N.; Wightman, R. M., Removal of differential capacitive interferences in fast-scan cyclic voltammetry. *Anal. Chem.* **2017,** *89* (11), 6166-6174. https://doi.org/10.1021/acs.analchem.7b01005
86. Johnson, J. A.; Rodeberg, N. T.; Wightman, R. M., Measurement of basal neurotransmitter levels using convolution-based nonfaradaic current removal. *Anal. Chem.* **2018,** *90* (12), 7181-7189. https://doi.org/10.1021/acs.analchem.7b04682
87. Yoshimi, K.; Weitemier, A., Temporal differentiation of pH-dependent capacitive current from dopamine. *Anal. Chem.* **2014,** *86* (17), 8576-8584. https://doi.org/10.1021/ac500706m
88. Tian, S.-Y.; Deng, S.-P.; Chen, Z.-X., Multifrequency large amplitude pulse voltammetry: A novel electrochemical method for electronic tongue. *Sens. Actuators, B* **2007,** *123* (2), 1049-1056. https://doi.org/10.1016/j.snb.2006.11.011
89. Takmakov, P.; Zachek, M. K.; Keithley, R. B.; Bucher, E. S.; McCarty, G. S.; Wightman, R. M., Characterization of local pH changes in brain using fast-scan cyclic voltammetry with carbon microelectrodes. *Anal. Chem.* **2010,** *82* (23), 9892-9900. https://doi.org/10.1021/ac102399n
90. Jo, T.; Yoshimi, K.; Takahashi, T.; Oyama, G.; Hattori, N., Dual use of rectangular and triangular waveforms in voltammetry using a carbon fiber microelectrode to differentiate

norepinephrine from dopamine. *J. Electroanal. Chem.* **2017,** *802*, 1-7. https://doi.org/10.1016/j.jelechem.2017.08.037
91. Zhu, F.; Yan, J.; Sun, C.; Zhang, X.; Mao, B., An electrochemical method for selective detection of dopamine by depleting ascorbic acid in diffusion layer. *J. Electroanal. Chem.* **2010,** *640* (1), 51-55. https://doi.org/10.1016/j.jelechem.2010.01.006
92. Kishida, K. T.; Saez, I.; Lohrenz, T.; Witcher, M. R.; Laxton, A. W.; Tatter, S. B.; White, J. P.; Ellis, T. L.; Phillips, P. E. M.; Montague, P. R., Subsecond dopamine fluctuations in human striatum encode superposed error signals about actual and counterfactual reward. *Proc. Natl. Acad. Sci. U.S.A.* **2016,** *113* (1), 200-205. https://doi.org/10.1073/pnas.1513619112
93. Meunier, C. J.; McCarty, G. S.; Sombers, L. A., Drift subtraction for fast-scan cyclic voltammetry using double-waveform partial-least-squares regression. *Anal. Chem.* **2019,** *91* (11), 7319-7327. https://doi.org/10.1021/acs.analchem.9b01083
94. Vasudevan, R. K.; Ziatdinov, M.; Vlcek, L.; Kalinin, S. V., Off-the-shelf deep learning is not enough, and requires parsimony, Bayesianity, and causality. *Npj Comput. Mater.* **2021,** *7* (1), 16. https://doi.org/10.1038/s41524-020-00487-0
95. Ziatdinov, M. A.; Ghosh, A.; Kalinin, S. V., Physics makes the difference: Bayesian optimization and active learning via augmented Gaussian process. *Mach. Learn.: Sci. Technol.* **2022,** *3* (1), 015003. https://doi.org/10.48550/arXiv.2108.10280
96. Jaworski, A.; Wikiel, H.; Wikiel, K., Benefiting from information-rich multi-frequency AC voltammetry coupled with chemometrics on the example of on-line monitoring of leveler component of electroplating bath. *Electroanalysis* *35* (1), e202200478. https://doi.org/10.1002/elan.202200478
97. Zhao, C.; Liu, Q.; Cheung, K. M.; Liu, W.; Yang, Q.; Xu, X.; Man, T.; Weiss, P. S.; Zhou, C.; Andrews, A. M., Narrower nanoribbon biosensors fabricated by chemical lift-off lithography show higher sensitivity. *ACS nano* **2021,** *15* (1), 904-915. https://doi.org/10.1021/acsnano.0c07503
98. Mitchell, E. C.; Dunaway, L. E.; McCarty, G. S.; Sombers, L. A., Spectroelectrochemical characterization of the dynamic carbon-fiber surface in response to electrochemical conditioning. *Langmuir* **2017,** *33* (32), 7838-7846. https://doi.org/10.1021/acs.langmuir.7b01443
99. Hibbert, D. B., Experimental design in chromatography: A tutorial review. *J. Chrom. B.* **2012,** *910*, 2-13. https://doi.org/10.1016/j.jchromb.2012.01.020
100. Https://scikit-optimize.Github.Io/stable/index.Html.
101. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V., Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011,** *12*, 2825-2830.