Using voltammetry augmented with physics-based modeling and Bayesian hypothesis testing to identify analytes in electrolyte solutions

Alexis M. Fenton Jr. and Fikile R. Brushett*

Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, U.S.A.
Email addresses: afenton@mit.edu (Alexis M. Fenton Jr.), brushett@mit.edu (Fikile R. Brushett)
**Corresponding Author: Fikile R. Brushett; 77 Massachusetts Avenue, 66-470, Cambridge, MA 02139*

Abstract

Voltammetry is a foundational electrochemical technique that can qualitatively and quantitatively probe electroactive species in solutions and as such has been used in numerous fields of study. Recently, automation has been introduced to extend the capabilities of voltammetric analysis through approaches such as Bayesian parameter estimation and compound identification. However, opportunities exist to enable more versatile methods across a wider range of solution compositions and experimental conditions. Here, we present a protocol that uses experimental voltammetry, physics-driven models, binary hypothesis testing, and Bayesian inference to enable robust labeling of analytes in multicomponent solutions across multiple techniques. We first describe the development of this protocol, and we subsequently validate the methodology in a case study involving five *N*-functionalized phenothiazine derivatives. In this analysis, the protocol correctly labeled solutions each containing 10H-phenothiazine and 10-methylphenothiazine from both cyclic voltammograms and cyclic square wave voltammograms, demonstrating the ability to identify redox-active constituents of a multicomponent solution. Finally, we identify areas of further improvement—such as achieving greater detection accuracy—and future applications to potentially enhance *in situ* or *operando* diagnostic workflows.

Keywords: Cyclic voltammetry, Cyclic square wave voltammetry, Analyte identification, Bayesian inference, Physics-based modeling, Phenothiazine

Published in the *Journal of Electroanalytical Chemistry* on January 1, 2022.

1. Introduction

Voltammetry is a foundational technique in electrochemical science that enables both qualitative and quantitative characterization of electroactive species—such as analytes (i.e., redox-active compounds)—for a variety of applications.[1–10] Examples include tracking the transient behavior of solutions in electrochemical systems[8,11–13] and labeling compounds within a sample.[14–17] When the composition of an analyte solution (i.e., the solvent, supporting salt, and redox-active species), along with the electrode surface morphology, are known prior to and remain constant throughout the experiment, established fundamental relationships can often be leveraged to discern both physical and electrochemical properties in the system of interest,[3,5,7,18] enabling mechanistic insights into electrochemical and chemical reactions of electroactive compounds.[19–24] However, in cases where the solution composition is unknown or evolves during the experiment, voltammetric analysis has primarily relied on qualitative visual methods, like peak appearance / disappearance and location.[14,15] Such approaches have practical utility,[14,15,17] but quantitative treatments are hampered because—unlike other techniques (e.g., nuclear magnetic resonance (NMR), Fourier transform infrared spectroscopy)—the output signal is not correlated to the molecular structure or connectivity.[20] Accordingly, additional analyses are typically required to identify constituent components, resulting in workflows that may be time-consuming and expensive.[1,8,11,25–29] In many cases, the standard techniques employed provide incomplete or misleading information, as necessary preparatory steps modify solution compositions, such as dilution with deuterated solvents for proton NMR analysis[11] or product purification *via* column chromatography for mass spectrometry studies.[25] In addition, it is not trivial to capture transient electrochemical processes with *ex situ* measurements.[27,28] Advances in voltammetric methods may enable streamlined component identification workflows by reducing the materials, equipment, and time intensity of

sample characterization, as these techniques can more readily probe analyte solutions in their native environment.[8,30]

To capitalize on the inherent advantages of voltammetry, black-box—that is, physics-agnostic—classification protocols for voltammetric labeling (i.e., identifying electroactive compounds from voltammograms) have been reported. These findings suggest that automation can be leveraged to characterize compounds in solutions by identifying and incorporating features that are challenging or impossible to ascertain from qualitative inspection.[29–34] While the prior literature shows promise, opportunities exist for further improvement. For example, protocols can identify compounds less accurately when the training and testing data are obtained under different supporting salt concentrations,[31] potentially necessitating a separate training dataset for each condition tested. Further, other approaches do not evaluate all species combinations in a multicomponent sample and thus may be less effective when a larger number of possible compositions is considered.[32,33] For example, Dean *et al*. labeled a solution containing a mixture of Cd, Hg, and Pb cations among other single-component samples,[32] but as the authors only considered this single combination, it is uncertain whether their protocol would retain accuracy if all combinations of the metal cations studied (e.g., Cd-Cu-Pb, Hg-Cu-Pb, etc.) were exhaustively considered. Finally, we note that regressive methods—such as partial least squares—may simultaneously estimate the concentrations of multiple constituents in a sample, which theoretically enables them to consider all possible species combinations, but their prediction error (either systemic or random) leaves ambiguity as to whether a compound is present at low concentrations or absent.[34] As such, there is a continued need to advance robust and flexible methods to identify compounds in solutions.

Existing approaches can be augmented to address these areas by incorporating physics-based modeling and by evaluating the presence of each compound individually using binary hypothesis testing[35]—i.e., each is present or absent. While the coupled reaction-transport phenomena that govern voltammetric responses are well-known,[5,19] these processes are not yet widely considered in automated voltammetric labeling methodologies, which may partially explain why prior protocols are challenged when evaluated at different analyte and supporting salt concentrations.[31,32] By incorporating physical phenomena into model formulations, voltammograms may be accurately simulated across a range of species concentrations using multiple experimental techniques with a single set of electrochemical and transport descriptors. Further, the number of possible compositions for a multicomponent solution scales combinatorically with the total number of species, potentially rendering exhaustive evaluation infeasible for larger sets. However, labeling scales linearly with the number of components if the presence of each compound is individually assessed to characterize the overall sample. Thus, automated labeling protocols may potentially enable analysis of more complex multicomponent systems.

Physical models and binary hypothesis testing can be readily adopted to identify compounds with Bayesian inference, which is used to infer the state of a system by recursively combining previous information of a process with new observations. While prior knowledge can come from arbitrary sources, Bayesian inference provides a quantitative basis to integrate information from previous observations, resulting in informed estimations; it is also an apt framework to conduct binary hypothesis testing.[35] With the voltammetric labeling presented in this work, the results for each compound can be combined to estimate the overall makeup, which, in turn, can be used to inform subsequent studies that may involve different techniques. Bayesian methods are also a

powerful framework to evaluate physics-informed models across multiple disciplines,[35–38] including electrochemistry;[39–43] for example, researchers have shown that parameters for electrochemical reactions can be reliably estimated using sensitive voltammetric techniques, mathematical optimization, and Bayesian inference.[44] However, to the best of our knowledge, Bayesian inference has not been used for voltammetric labeling.

As such, this work seeks to develop a simple yet versatile protocol that builds on previous work by combining experimental voltammetric data, physical modeling, binary hypothesis testing, and Bayesian inference to simultaneously identify multiple analytes in a solution. This protocol is validated with a case study involving a set of five phenothiazine derivatives and two voltammetry techniques: cyclic voltammetry (CV) and cyclic square wave (CSW) voltammetry. Specifically, we demonstrate that the protocol can differentiate between analytes even when the testing and training datasets are obtained with different experimental techniques. Consequently, this methodology may reduce the training data needed to probe samples across different solution compositions, environmental conditions, and experimental techniques, enabling accelerated *in situ* or *operando* labeling. This protocol can also enable fewer and more targeted follow-up *ex situ* techniques that may be integrated with previous estimations through the Bayesian framework, reducing the time and resources needed to fully characterize more diverse samples. To enable further development and expansion of this protocol, we provide MATLAB® code on our GitHub page (https://github.com/afentonjr/BayES-Lab) that constructs a compound library from CSW voltammograms and, using that library, labels analytes from previously unexamined experimental data.

2. Methods

2.1 Overview

To accurately label analytes with voltammetry, a library cataloguing compounds must first be developed and subsequently proven effective in the identification of analytes from previously unseen experimental data. Accordingly, the experimental acquisition methods and the library development are respectively detailed in **Sections 2.2** and **2.3**. Once a high-fidelity library is fully constructed, it can be fit to new experimental data to label analytes. This labeling process, in turn, involves two major steps. The first (**Section 2.4**) is a regressive step that fits catalogued species to an experimental voltammogram, where all analytes in a library are fit to yield a vector of best-fit concentrations; each vector entry corresponds to a single catalogued compound. The second step (**Section 2.5**) involves using the same experimental dataset to label the multicomponent system being probed; in this instance, each analyte in the library is evaluated to determine its probability of being present using binary hypothesis testing and Bayesian inference. To increase accessibility and degree of implementation, the code used for both library construction and analyte identification, along with subroutines, are available on GitHub: https://github.com/afentonjr/BayES-Lab.

2.2 Experimental

All chemicals were used as received, and all experiments were conducted in a glovebox (MBraun Labmaster, H_2O < 5 ppm, O_2 < 1 ppm) filled with argon (Airgas, purity of *ca*. 100 %, catalog number AR UHP300). The glovebox temperature was measured to be 25.5 °C and 25 °C on two occasions using a glass thermometer (VWR®, ± 2 °C). All the phenothiazines, the tetrabutylammonium hexafluorophosphate ($TBAPF_6$, Sigma Aldrich, ⩾ 99 %, 86879), and the

dichloromethane (ACROS Organics^{TM}, 99.9 %, AC610931000) were opened and stored in the glovebox. All materials were directly added from their container to a 10 mL volumetric flask with a plastic spatula to ensure the mass of material in the solution matched the balance reading (Mettler Toledo, Balance XS64, 61 g capacity with ±0.1 mg readability). Every solution studied contained 1 mM of each analyte, along with 0.1 M $TBAPF_6$ in dichloromethane. Ferrocene (Sigma Aldrich, 98 %, F408) was used as an internal standard for the reference electrode[45] at a concentration of *ca*. 1 mM. The working electrode was a glassy carbon disk electrode (CH Instruments, 3 mm dia., CHI104) polished with 0.05 μm alumina powder (Buehler MicroPolish Powder, 4010075) in deionized water (18.2 MΩ cm). The reference electrode was either a Ag/Ag^+ electrode using a non-aqueous reference electrode kit (MF-2062) filled with 0.1 M $AgPF_6$ (Sigma Aldrich, 98 %, 208361) in acetonitrile (Fisher, Certified ACS, A21-1) or, when the first reference electrode was unavailable, an aqueous Ag/AgCl (3 M NaCl) electrode (BASi, MF-2052) brought into the glovebox and stored in vial containing propylene carbonate (Gotion, 99.99 %) without any supporting salt during experiments. The counter electrode, in turn, was a Pt coil electrode (BASi, 99.95 %, MW-1033). When not in use, the Ag/Ag^+ reference was stored in the glovebox in the same fill solution, and the Ag/AgCl reference was stored outside the glovebox in a solution of 1 M KCl.

Five phenothiazines (**Figure 3**)—synthesized and purified as described in the SI by the Odom Research Group at the University of Kentucky—were catalogued to create the library used for validating the labeling protocol (**Section 3**): 10H-phenothiazine (i.e., unsubstituted phenothiazine, PT), 10-methylphenothiazine (MPT), 10-ethylphenothiazine (EPT), 10-isopropylphenothiazine (iPrPT), and 10-phenylphenothiazine (PhPT). Two separate solutions, both containing only a single phenothiazine at a concentration of 1 mM, along with 0.1 M $TBAPF_6$ in dichloromethane,

were examined to estimate the electrochemical and transport parameters for the corresponding analyte in library development. Three solutions each containing a mixture of 1 mM PT and 1 mM MPT (also with 0.1 M $TBAPF_6$ in dichloromethane) were used to test the identification protocol.

For each prepared solution, two voltammetry techniques were conducted—CV and CSW voltammetry—using either a VMP-3 potentiostat (BioLogic) or a VMP-300 potentiostat (BioLogic) using EC-Lab® software and processed with Microsoft Excel and MATLAB® R2020a. Cyclic voltammograms were obtained at 25, 50, 100, 200, 500, and 1000 mV s^{-1}, with all voltammograms corrected for resistance-driven potential distortions using the BioLogic protocol "*iR* determination with electrochemical impedance spectroscopy" (the "ZIR" protocol).[46] For the "ZIR" protocol, the working electrode potential was set to its open-circuit value. A sinusoidal potential with a 20 mV amplitude and a 100 kHz frequency was applied, a delay of 10 % of the period duration was added before the measurement, and the reported resistance was averaged over four measurements. The resistance was compensated either 100 % or 85 % by the software during the experiment, with the remaining percentage manually post-corrected; in some cases, the solution resistance was not fully compensated during acquisition to avoid possible oscillations in the potentiostat.[46] For all voltammetry experiments, the bandwidth was manually adjusted *via* trial and error to minimize noise in the current acquisition. The potential bounds varied for each analyte; the most negative and initial potential was set to be approximately 400-500 mV negative of the ferrocene redox potential. The most positive (and the turnaround) potential of the voltammetric experiment, in turn, was set to be between 200-400 mV positive of the redox potential of the phenothiazine(s) probed. More specifically, the most positive potential was set far enough away from the phenothiazine redox potential as to minimally influence the voltammogram shape[3] but not so far as to access the second electron transfer event of the phenothiazine to a considerable

extent[8] or to oxidatively decompose the solution or the electrode. Generally, the upper bound was found *via* visual inspection using CV at a 50 mV s^{-1} scan rate. After each cyclic voltammogram was obtained, no electrochemical experiments were conducted for either 5 min (50-1000 mV s^{-1}) or 10 min (25 mV s^{-1}) before the next to allow the boundary layer to reset.

CSW voltammograms were obtained using the same potential bounds as those in the cyclic voltammograms. The step height was 10 mV, the pulse height was 50 mV, and the pulse duration (per half-period) was 100 ms, resulting in an effective scan rate of 50 mV s^{-1}. The potential was held at the initial, most negative (i.e., reductive) potential for 2 s before the initial positive (oxidizing) sweep, and the reported current for each potential step was calculated by averaging the raw current over the last 30 % of the step. Six CSW voltammograms were obtained at these same conditions for each solution tested. The "ZIR" protocol was performed the same way as that with CV, and each CSW voltammetry experiment was separated by a 5 min wait. Following the initial suite of CV and CSW voltammetry tests, 1 mM of ferrocene was added, and the experiments were repeated to calibrate the potential axis to that of a known redox event.[45] For library development, two separate solutions of every phenothiazine (five phenothiazines, so 10 solutions in total) were tested, each containing approximately the same analyte concentration of 1 mM. This procedure resulted in 12 CSW voltammetry datasets with the same potential waveform and two CV datasets for all six scan rates for each phenothiazine; only the CSW voltammograms were used to construct the library. For protocol validation, three solutions of the phenothiazine mixture were examined, resulting in 18 CSW voltammograms acquired using the same potential waveform and three cyclic voltammograms at six different scan rates.

2.3 Library development

In this work, library development involves generating a characteristic set of electrochemical and transport descriptors (*vide infra*) for each analyte by first acquiring experimental data, subsequently simulating modeled voltammograms, and then comparing the two using both weighted least squares regression and Bayesian inference. We note that this approach is not the only viable method, but it does possess favorable properties as compared to other options. For example, literature data mining[47] could be used, but the natural language processing necessary to implement this method is non-trivial.

2.3.1 Experimental data acquisition

We elected to use CSW voltammetry to acquire the data for library construction because it can be more accurately modeled; its waveform minimizes background electrochemical signals while amplifying Faradaic processes,[20,42,48] offering an advantage over CV.[4,49] For this reason, square wave (SW) voltammetry—that is, CSW voltammetry without the reverse sweep—has been employed in many studies involving qualitative analyses.[14,15,17] Further, CSW voltammetry can more readily discern various electron transfer mechanisms (e.g., an electron transfer followed by the homogeneous degradation of the product[50]) than SW voltammetry by virtue of the reverse sweep. We also conduct repeats of experiments for statistical rigor and to calculate the experimental standard deviation (*vide infra*). As mentioned in **Section 2.2**, each analyte studied in this work was catalogued using data from two solutions; for each, six CSW voltammograms were acquired using the same potential waveform, resulting in 12 total experimental CSW voltammograms for each analyte in the library.

2.3.2 Model development

The theoretical models used in this work are necessary both for parameter extraction and analyte labeling and as such are discussed here. For each compound in the library, two models (diffusion rate-limited and kinetic rate-limited one-electron transfers) were simulated using the one-dimensional transient diffusion equation with an electrochemical reaction on a planar, impermeable, and ideal (non-fouling) electrode surface. The reaction considered is **Equation (1)**.

$$R \rightleftharpoons O + e^- \tag{1}$$

In **Equation (1)**, a phenothiazine (R) oxidizes to a radical cation (O) in a one-electron transfer; the mass conservation equations are expressed in **Equation (2)**. Note that all values are non-dimensional unless otherwise noted.

$$\begin{aligned} \frac{\partial c_R}{\partial \tau} &= \frac{\partial^2 c_R}{\partial \xi^2} \\ \frac{\partial c_O}{\partial \tau} &= d_O \frac{\partial^2 c_O}{\partial \xi^2} \end{aligned} \tag{2}$$

In **Equation (2)**, lower-case c_i is the dimensionless concentration of species i ($c_i = C_i \cdot C_{R,bulk}^{-1}$, where upper-case C_i is the dimensional concentration and $C_{R,bulk}$ is the dimensional concentration of R far from the electrode, both in units of mol m^{-3}), $\tau = tD_R \cdot r_e^{-2}$ is dimensionless time (D_i is the diffusion coefficient of species i in units of m^2 s^{-1}, t is dimensional time in units of s, and r_e is the electrode radius in units of m), $\xi = x \cdot r_e^{-1}$ is the dimensionless length (x is the distance from the electrode in units of m), and d_O is the ratio of diffusion coefficients $D_O \cdot D_R^{-1}$.

These coupled dimensionless differential equations are subject to the boundary and initial conditions expressed in **Equations (3) - (6)** for a single electron transfer. **Equations (5)** and **(6)** are mutually exclusive and should not be simultaneously used; **Equation (5)** is used for a diffusion

rate-limited one-electron transfer, while **Equation (6)** is used for an kinetic rate-limited one-electron transfer.[19]

$$\begin{aligned} c_R(\xi \to \infty, \tau) &= c_R(\xi, \tau = 0) = 1 \\ c_O(\xi \to \infty, \tau) &= c_O(\xi, \tau = 0) = 0 \end{aligned} \tag{3}$$

$$\left.\frac{\partial c_R}{\partial \xi}\right|_{\xi=0} + d_O \left.\frac{\partial c_O}{\partial \xi}\right|_{\xi=0} = 0 \tag{4}$$

$$c_R(\xi = 0, \tau) = c_O(\xi = 0, \tau) \exp\left(\frac{-F\eta(\tau)}{R_G T}\right) \tag{5}$$

$$d_O \left.\frac{\partial c_O}{\partial \xi}\right|_{\xi=0} = K_0 \left[c_O(\xi = 0, \tau) \exp\left(\frac{-\alpha F\eta(\tau)}{R_G T}\right) - c_R(\xi = 0, \tau) \exp\left(\frac{(1-\alpha) F\eta(\tau)}{R_G T}\right) \right] \tag{6}$$

In **Equations (3)-(6)**, $\eta = E - E_0$ is the overpotential in units of V (E is the applied potential and E_0 is the formal potential of the redox couple, both in units of V vs. a reference redox event), R_G is the universal gas constant (8.314 J mol^{-1} K^{-1}), T is the absolute temperature in units of K (set to 298.15 K based on the measured glovebox temperatures), F is the Faraday constant (96485 C mol^{-1}), $K_0 = k_0 r_e \cdot D_R^{-1}$ is the dimensionless heterogeneous rate constant (k_0 is the dimensional analog in units of m s^{-1}), and α is the transfer coefficient (dimensionless). **Equation (3)** assumes that only species R is present before the experiment and far away from the working electrode at all times, while **Equation (4)** relates the flux of both species at the electrode surface. **Equation (5)** relates the surface concentration of the species *via* the Nernst equation (diffusion rate-limited electron transfer), while **Equation (6)** relates the surface flux of O to the surface concentrations *via* the Butler-Volmer relation (kinetic rate-limited electron transfer).

These voltammograms were numerically simulated; details on the implementation scheme are in the SI (**Section S.2.4**) based on an established framework.[19] Briefly, each voltammogram was

simulated with a discretization of 20000 steps in potential per unit volt (i.e., 5×10^{-5} V per step) using either MATLAB® R2020a on an Intel® Core™ i7-7500U CPU @ 2.70 GHz 2.90 GHz laptop computer or MATLAB® (either R2019b or R2020a) on the MIT Supercloud supercomputing resource;[51] the former computing resource took *ca.* 8 s to simulate each voltammogram. Note that, as the simulation time scales approximately linearly with total number of discretization steps, the speed can be increased by introducing coarser discretization albeit at the expense of accuracy.

2.3.3 Fitting procedure

The experimental and simulated voltammograms generated according to the procedures outlined in **Sections 2.3.1** and **2.3.2** are then compared to perform parameter estimation. These were fit by simultaneously adjusting the values of all the electrochemical and transport descriptors. Specifically, E_0, D_R, and D_O (for the diffusion rate-limited electron transfer mechanism) or E_0, D_R, D_O, k_0, and α (for kinetic rate-limited electron transfers) were introduced to the simulator as adjustable parameters—whose values were bounded based on either convention or observations from experimental data—to find which parameter set, designated as the vector $\boldsymbol{\theta}$, maximized the model likelihood. The likelihood function (**Equation (7)**) is assumed to be a product of the probability distribution functions (PDFs) of the error at each overpotential, which are assumed to be normal and independent of each other (i.e., the errors are random and not systemic nor reliant on errors at other overpotentials).

$$f\left(\boldsymbol{I}_{exp};\boldsymbol{I}_m(\boldsymbol{\eta},\boldsymbol{\theta}),\boldsymbol{\sigma}\right)=\prod_{j=1}^{N}\frac{1}{\sigma_j\sqrt{2\pi}}\exp\left(-\frac{\left(I_{exp,j}-I_{m,j}(\eta_j,\boldsymbol{\theta})\right)^2}{2\sigma_j^2}\right) \tag{7}$$

In **Equation (7)**, $f\left(\boldsymbol{I}_{exp};\boldsymbol{I}_{m}\left(\eta,\theta\right),\sigma\right)$ is the PDF of the experimentally measured current $\boldsymbol{I}_{exp}$ (treated as a random variable) parameterized by the modeled current $\boldsymbol{I}_{m}$ and the experimental standard deviation σ (all in units of A); as such, $f\left(\boldsymbol{I}_{exp};\boldsymbol{I}_{m}\left(\eta,\theta\right),\sigma\right)$ is equivalent to the PDF of the error (i.e., the difference between experimental and modeled currents). $\boldsymbol{I}_{m}$ is a deterministic function of the overpotential vector $\boldsymbol{\eta}$ and parameter vector $\boldsymbol{\theta}$; note that difference currents are often used in the place of absolute currents for CSW voltammetry.[5,20] $\boldsymbol{\sigma}$, in turn, was either calculated between the 12 CSW voltammograms (for library construction) or estimated (for analyte identification). Bolded terms are vector quantities containing all N data points in the voltammogram, and the index j refers to the j^{th} entry of each vector. As a result, maximizing **Equation (7)** is equivalent to minimizing the sum of the magnitude of the exponential arguments by adjusting $\boldsymbol{\theta}$ for either electron transfer model; namely, $\hat{\boldsymbol{\theta}}=\arg\max_{\theta} f=\arg\min_{\theta}\sum_{j=1}^{N}\left(\left(I_{exp,j}-I_{m,j}(\eta_j,\boldsymbol{\theta})\right)\sigma_j^{-1}\right)^2$, where $\hat{\boldsymbol{\theta}}$ is the optimal set of parameters. The most likely electron transfer mechanism was then chosen using binary hypothesis testing and Bayesian inference (*vide infra*); this mechanism was then catalogued as a descriptor for the analyte being assessed. Optimal parameters corresponding to the selected electron transfer mechanism were subsequently recorded as the remaining descriptors, all of which are reported in the SI (**Table S3**). We note that the reported optimal parameter set may be nonunique,[3,52] but degenerate sets describe similar voltammetric curves, meaning the uniqueness of the parameters is not expected to impact the ability of the protocol to differentiate between analytes. The construction of each library entry (i.e., individual analytes) was performed on the MIT Supercloud supercomputing resource[51] using 80 or 100 cores, taking 7-10 days to complete. We note that less intensive computational resources (e.g., the local computing resource listed above) may be able to output

sufficient—although perhaps not as accurate—descriptors in a shorter time frame (*ca*. 1 h) by using a coarser time mesh and fewer initial guesses. We do not anticipate the predictive power of the library to be adversely affected by such a change provided the same mesh is used throughout the entire process.

2.4 Library-data fitting

Once a library is constructed, it can be applied to new experimental data—specifically, 18 CSW voltammograms and 18 cyclic voltammograms (as mentioned in **Section 2.2**)—to estimate how much of each species is present and, ultimately, to label analytes in solutions. To achieve this, the information from the library is first combined with the same input waveform used to acquire the new experimental dataset to simulate a concentration-normalized current for each analyte (**Figure 1**). These normalized simulated voltammograms are then regressed to the experimental data by adjusting the concentration weights to maximize the likelihood function (**Equation (7)**), which is equivalent to the boxed optimization in **Figure 1b**. This fitting procedure yields a vector of best-fit concentrations for all the library constituents, where each vector entry estimates the concentration of the corresponding compound. However, this vector can include analytes that are not actually present in the sample due, in part, to random errors whose effects are challenging to physically quantify (e.g., random heterogeneities generated from electrode polishing[39]). Consequently, it is necessary to evaluate the inclusion of every analyte by assigning to each a probability of existence.

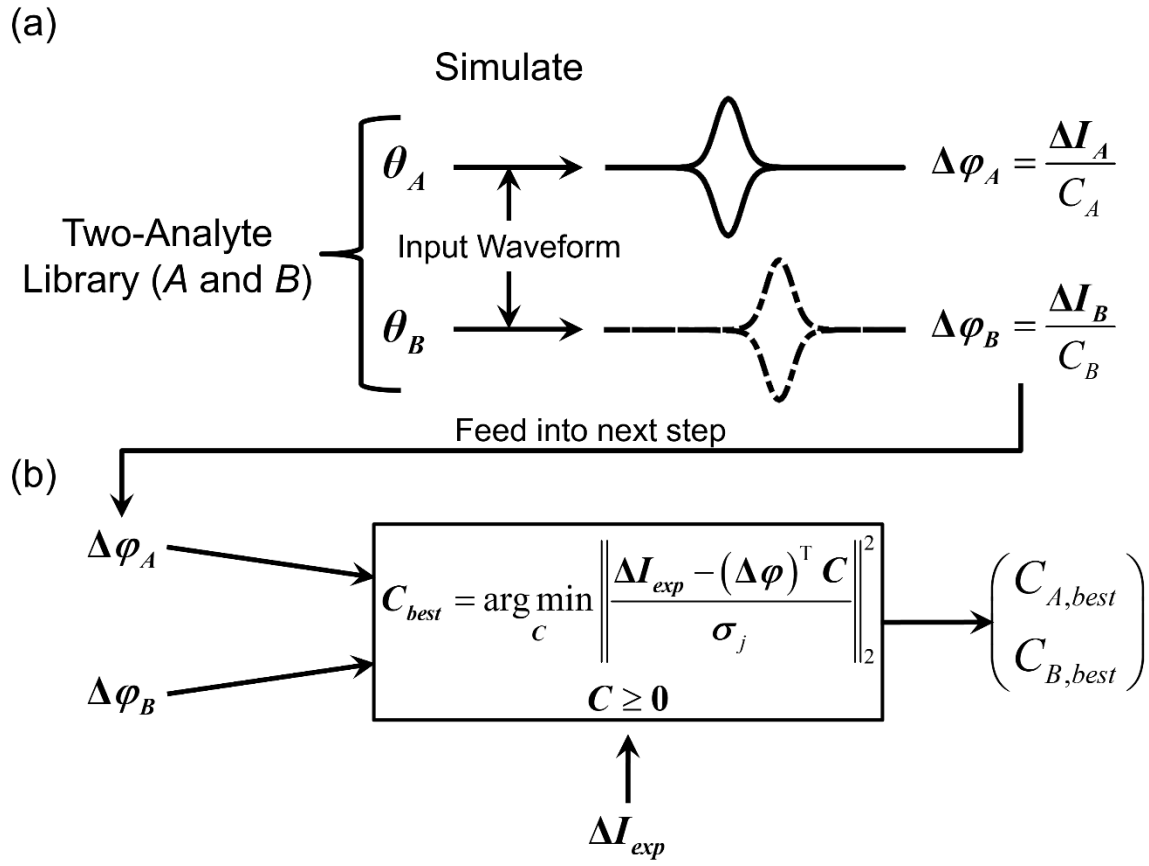

Figure 1. Procedure used to estimate the concentration of each catalogued library compound in an experimental dataset using CSW voltammetry as an example technique. (a) An example two-analyte library is used to simulate the concentration-normalized difference current ($\Delta\varphi_i;\ i = A, B$) using the catalogued list of descriptors (θ_i) and the same input waveform used to acquire the experimental dataset. (b) The resulting concentration-normalized difference current is compared with the experimental data ($\Delta \boldsymbol{I}_{exp}$) using weighted linear least squares fitting to estimate the concentrations, $C_{A,best}$ and $C_{B,best}$, in the sample. The concentration-normalized difference current is linearly proportional to the bulk concentration, enabling rapid optimization.

2.5 Analyte identification

Once the simulated voltammograms are fit to experimental data, candidate analytes are culled by evaluating models for two hypotheses (H_0 and H_1): the null hypothesis, in which all analytes except the one of interest are included ("exclusion", H_0), and the alternative hypothesis, where all library entries—including the analyte of interest—are considered ("inclusion", H_1). We note that this framework is also used to estimate the electron transfer mechanism of an analyte during library development (**Section 2.3**); there, H_0 and H_1 respectively represent a diffusion and kinetic rate-limited electron transfer (or *vice versa*). The probabilities for these hypotheses are calculated using Bayesian inference; the hypothesis with a probability of greater than 50 % is the accepted state based on the Maximum *a Posteriori* probability (MAP) rule.[35] This process is detailed in **Equations (8)** and **(9)** (for further details, see **Equations S1-S4**).

$$P(H_i \mid O_{obs}) = \frac{f(O_{obs} \mid H_i)P(H_i)}{f(O_{obs})} \tag{8}$$

In **Equation (8)** (Bayes' Rule), P is a discrete probability, f represents a PDF, H_i is the i^{th} hypothesis, and O_{obs} is an observation. $P(H_i)$ is the prior probability (i.e., the probability that H_i is true before the observation was made)—assumed to be equal for all hypotheses (50 %) in this work— $P(H_i \mid O_{obs})$ is the posterior probability (the probability that H_i is true given the observation), and $f(O_{obs} \mid H_i)$ is the likelihood PDF of observing O_{obs} given that H_i is true. $f(O_{obs})$ is the PDF of O_{obs} occurring across all hypotheses considered. We note that $f(O_{obs}) = \sum_{q=1}^{M} f(O_{obs} \mid H_q)P(H_q)$, where q is a counter for M total hypotheses.

In the context of this work, $P(H_i)$ is the probability that hypothesis i (either the presence of a single analyte or an electron transfer model) is true before evaluating experimental data, $f(O_{obs} \mid H_i)$ is a continuous PDF that evaluates for both the goodness of fit and the number of model parameters to prevent overfitting, and $P(H_i \mid O_{obs})$ is the probability of hypothesis i being true after considering the experimental data.[35]

In analyte identification, there is also a possibility that the measured currents arise from background non-faradaic processes independent of the presence of any redox-active compounds. As such, the probability of a peak resulting from background noise (expressed as $P_{background}$) must also be evaluated to yield the final probability (**Equation (9)**).

$$P_f(H_i \mid O_{obs}) = P(H_i \mid O_{obs}) \times (1 - P_{background}) \tag{9}$$

In **Equation (9)**, $P_f(H_i \mid O_{obs})$ is the final probability reported for a given analyte. Note that this formula assumes that $P(H_i \mid O_{obs})$ and $P_{background}$ are independent of each other. This assumption is reasonable, as faradaic events are expected to influence background processes to a negligible extent. However, there theoretically may be instances where the presence of an analyte significantly impacts the behavior of the background current; such dependencies are not captured in scans of an electrolyte solution (only supporting salt and solvent) and thus are not captured by **Equation (9)**. In this work, the probability that an identified peak resulted from background processes was nearly zero (all the probabilities were zero within the working precision of MATLAB® R2020a). Nevertheless, this feature may become important in future scenarios, such as samples containing analytes at μM or nM concentrations.

The probability for each analyte is evaluated individually *via* binary hypothesis testing according to **Equations (8)** and **(9)**. Analytes with probabilities greater than 50 % are assumed to

be present—in line with the MAP rule—and *vice versa*; the workflow for this process is illustrated in **Figure 2**.

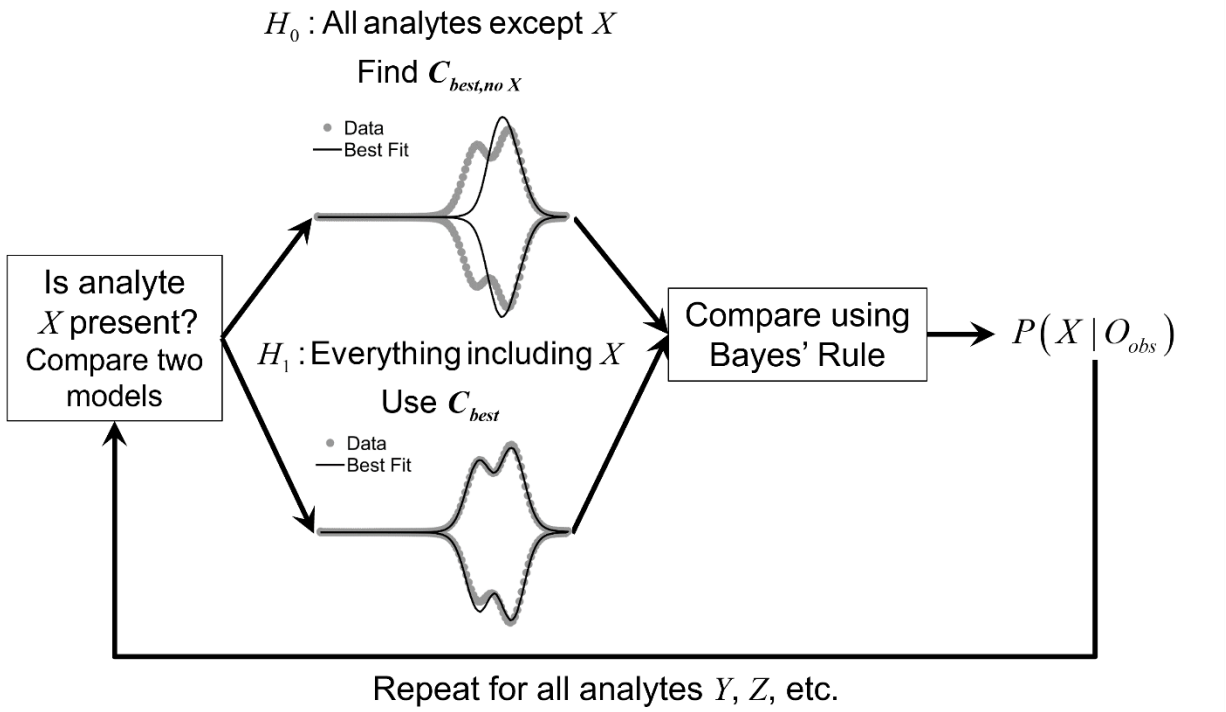

Figure 2. Schematic of procedure to assign probabilities for all analytes in the library one at a time *via* binary hypothesis testing. X is the analyte of interest; H_0 and H_1 refer to null (exclusion) and alternative (inclusion) hypotheses, respectively; and O_{obs} is the experimental observation (i.e., the experimental voltammogram).

Within the workflow depicted in **Figure 2**, the first model (H_0) represents the exclusion of the analyte in question. In it, CSW voltammograms for every catalogued analyte except the species of interest are fit to a single experimental dataset. The second model (H_1) represents the inclusion of

the analyte in question. In it, every analyte is considered; the resulting fit has already been evaluated when finding the vector of best-fit concentrations (**Figure 1**). These two models are then compared, and the probability of existence is determined using **Equations (8)** and **(9)**. This procedure is repeated for every analyte under consideration, and the resulting probabilities can then be assessed to determine which analytes are present. The process outlined in **Sections 2.4** and **2.5** take *ca*. 1 min to complete using two cores in MATLAB® R2020a with the local computational resource previously mentioned. This process is considerably faster than library construction because only a single set of voltammograms is simulated to fit models and data using the linear relationship between concentration and current observed in this work (**Section 2.4**). In comparison, library development requires many sets of voltammograms to be simulated because of the highly non-linear relationship between the current and relevant electrochemical parameters (**Section 2.3**).

3. Results and Discussion

3.1 Case study description

This protocol was validated with a case study involving phenothiazines, a class of redox-active organic compounds used for overcharge protection in Li-ion batteries[25,26,53] and, more recently, as analytes for the positive half-cell in redox flow batteries.[8,54] For this study, five different *N*-functionalized phenothiazine derivatives (PT, MPT, EPT, iPrPT, and PhPT), whose structures are depicted in **Figure 3**, were synthesized. Importantly, these compounds are stable in their neutral and singly-charged forms on the CV time scale (*ca.* 1 min).[25] We restrict ourselves to a single core for simplicity; different molecular classes are anticipated to be more easily differentiable based on variations in exhibited properties. We also anticipate that for more extensive applications (e.g., samples without entirely deterministic preparation), additional consideration will be needed to

create a manageable library that is not too large by vetting candidate compounds using intuition (e.g., excluding infeasible or unlikely species) and using *a priori* observation, such as eliminating a compound from contention if no voltammetric peak is recorded at its predicted redox potential. We also note that judicious library selection is important, as the output probabilities for each species are dependent on the compounds present in the library.

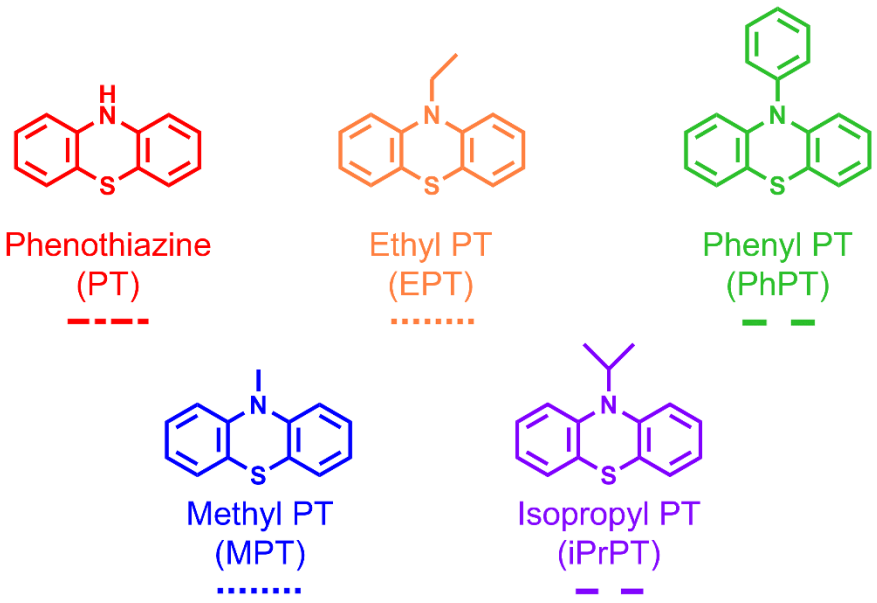

Figure 3. Structures and abbreviations of phenothiazines used as analytes in this study. Dashed, dotted, or dash-dotted lines indicate the line style used to plot data pertaining to each phenothiazine.

First, the phenothiazine library was constructed according to the procedure in **Section 2.3**; the generation process and library contents are described in detail within the SI (see **Section S.2**). This library was then applied to three identical solutions of known composition (containing 1 mM PT, 1 mM MPT, and 0.1 M $TBAPF_6$, all in dichloromethane) probed using both CSW voltammetry

and CV. Dichloromethane, used previously for electroanalytical studies,[25] is a non-nucleophilic solvent and is naturally anhydrous, creating favorable conditions for phenothiazine stability.[55] Additionally, we assumed that, under dilute conditions, the dissolved phenothiazines do not interact with each other during the electrochemical experiments, and thus, the voltammetric response was a superposition of the two individual species,[5] as illustrated in **Figure S3**.

Note that the methodology outlined in **Figure 2** allows for the identification of all 32 possible analyte combinations by evaluating each phenothiazine individually; as such, a successful case study will demonstrate that this protocol can deconvolute voltammograms comprised of multiple analytes while exhaustively considering all possible combinations. We note that while the experimenter knew the composition of the solution (ground truth), the protocol had no knowledge of the sample makeup prior to evaluating the experimental dataset; the routine was only offered the phenothiazine library, experimental data, and additional parameters not linked to analyte identities, such as working electrode radius, voltammetric waveform parameters, etc.

3.2 Protocol validation

After its construction, we used the phenothiazine library to simulate concentration-normalized voltammograms for each derivative, shown for CSW voltammograms in **Figure 4a**. These were fit to the experimental data to yield a vector of concentrations that best fit the data. **Figure 4b** illustrates the data for a representative experimental trial (one of 18 CSW voltammetry trials), the corresponding best-fit voltammogram, and the resulting concentration estimates.

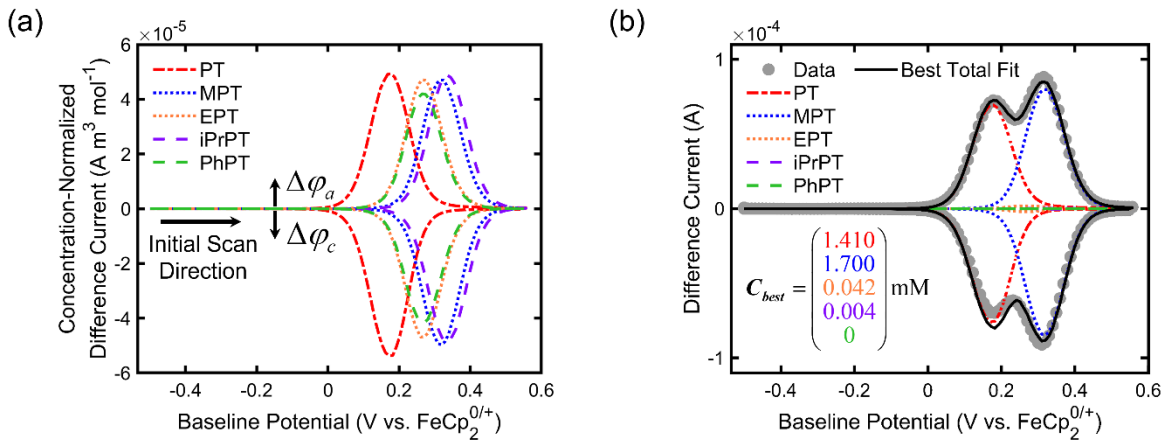

Figure 4. Generation and fitting of concentration-normalized CSW voltammograms to a single experimental dataset. (a) Concentration-normalized CSW voltammograms of the phenothiazines, extracted from the library used in this study. Anodic / oxidative currents (denoted by the subscript a) are positive in sign, cathodic / reductive (denoted by the subscript c) are negative in sign, and the initial potential sweep is from negative to positive potentials. This convention holds for all voltammograms depicted in this work. (b) Contribution of each phenothiazine to the best total fit of the experimental data and the corresponding best-fit concentrations (listed in the same order and color scheme as the legend). Note that difference currents are often reported in the place of absolute currents for CSW voltammetry.[5,20]

The vector of best-fit concentrations (C_{best}) in **Figure 4b** contained the estimated concentrations of each analyte in solution for a single trial. Across six CSW voltammetry and three CV trials from a single solution (nine total trials), the concentration estimated for unsubstituted PT had a 9.52 % error (**Equation S13**) with a standard deviation of 7.6×10^{-2} mM (**Equation S14**), and the concentration of MPT had a 8.06 % error with a standard deviation of 1.8×10^{-1} mM—

further details and discussion are found in the SI (**Section S.3**). We also note that the estimated concentrations of EPT and iPrPT were greater than zero, even though neither were present in reality. Specifically, in **Figure 4b**, the overall best fit was achieved by including EPT and iPrPT at estimated concentrations of 0.042 mM and 0.004 mM, respectively. However, in all nine cases, the second inferential step successfully excluded compounds not actually present from consideration (*vide infra*).

Once the best-fit vector of concentrations was estimated for a single dataset (**Figure 4b**), the probability of each phenothiazine being present was calculated according to the procedure outlined in **Figure 2**. For each phenothiazine studied, two models were examined: one considering every phenothiazine except the one currently being examined (representing exclusion— H_0), and another considering all five phenothiazines in the library (representing inclusion of the interrogated phenothiazine— H_1). The results from this analysis are depicted in **Figure 5**.

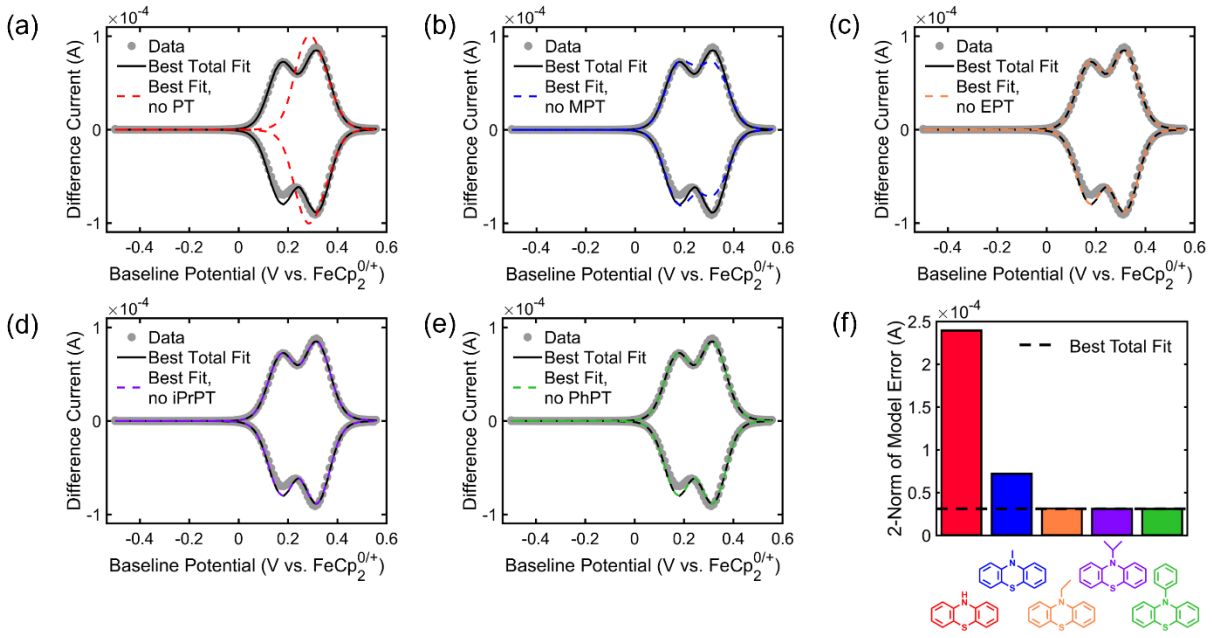

Figure 5. Graphical illustration of the best fits used in the labeling workflow when (a) PT, (b) MPT, (c) EPT, (d) iPrPT, and (e) PhPT are excluded from consideration (H_0, dashed lines) and when all phenothiazines are considered in the library (H_1, black line). Part (f) shows the 2-norm of the errors when each phenothiazine is excluded from consideration. Note that in parts (c)-(e), the dashed and solid lines are both present but are nearly or fully overlapping.

In **Figure 5**, the exclusion model fit (H_0, dashed lines) estimates whether an analyte of interest is present. If its fit is poorer than the inclusion model (H_1, black line), then the species is likely to be present because its inclusion is necessary to better fit the experimental data, and *vice versa*. From this, our preliminary conclusion—which will be evaluated more rigorously—is the protocol should label PT and MPT as present and EPT, iPrPT, and PhPT as absent. **Figure 5f** depicts the

2-norm error (**Equation (10)**) when phenothiazines are excluded from consideration to quantify the illustrations in **Figures 5a-e**.

$$\text{2-norm error (A)} = \sqrt{\sum_{j=1}^{N}\left(I_{exp,j} - I_{m,j}\right)^2} \tag{10}$$

The 2-norm error will also show that the protocol avoids overfitting by selecting models with fewer parameters (in this case, the exclusion model) if the error does not significantly increase compared to the inclusion model. Quantitative metrics on this balance between model simplicity and error are further discussed in the SI (**Section S.1**).

To substantiate these preliminary conclusions, the probabilities of each phenothiazine being present were calculated using **Equations (8)** and **(9)**. To demonstrate the repeatability and the adaptability of this protocol across different techniques, this procedure was applied to 18 CSW voltammetry datasets (all acquired with the same input waveform) and nine CV datasets (three at 25 mV s^{-1}, three at 50 mV s^{-1}, and three at 100 mV s^{-1}) across three solutions that were independently prepared. The probabilities were estimated for all 27 datasets, with the results depicted in **Figure 6**. To illustrate the worst-case scenario, the smallest probabilities for the sets of PT and MPT are reported, while the largest are reported for EPT, iPrPT, and PhPT.

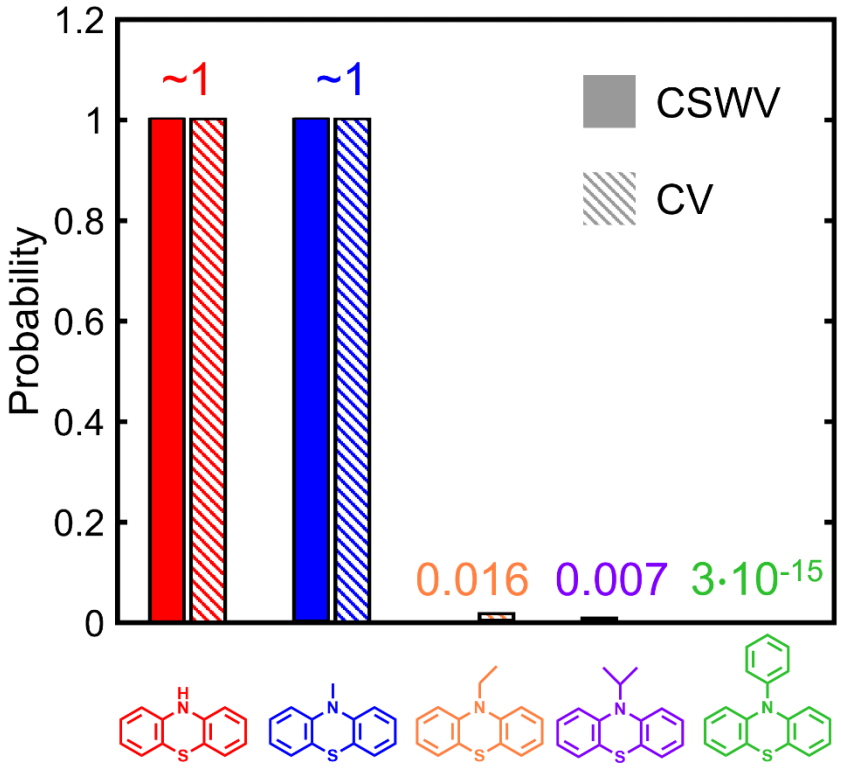

Figure 6. Bar graph depicting the probability of each phenothiazine residing in the solution being examined, in agreement with the phenothiazines known *a priori* to be present. To illustrate the worst-case scenario, the lowest probabilities for PT and MPT are plotted for both CSW voltammetry (written as "CSWV" in the figure) and CV data. Similarly, for EPT, iPrPT, and PhPT, the largest probabilities are plotted for both techniques. The reported value is the smaller of the two plotted probabilities for PT and MPT, and conversely, the larger of the two plotted probabilities for EPT, iPrPT, and PhPT.

Higher probabilities indicate that the phenothiazine of interest is more likely to be present, whereas the opposite is truth for lower probabilities. As expected from inspection of **Figure 5**, this

methodology found that for all 27 data sets considered, both PT and MPT were in solution while the other phenothiazines were not. Further, the accuracy and precision of the final concentration estimates (i.e., after culling the library) decreased and increased, respectively. Across the same nine trials previously mentioned, PT exhibited an error of 11.86 % (2.34 % increase) with a standard deviation of 4.3×10^{-2} mM (3.3×10^{-2} mM decrease), while MPT exhibited an error of 10.27 % (2.21 % increase) with a standard deviation of 1.6×10^{-1} mM (2.6×10^{-2} mM decrease). As such, the protocol both determines the identities and estimates the concentrations of the phenothiazines in the probed samples.

3.3 Discussion

This case study demonstrates that our methodology can identify multiple analytes and their estimated concentrations in solution using different voltammetric techniques. However, based on the relative positions of the phenothiazines in potential space, the observant experimentalist may conclude that PT is in solution *via* visual inspection, as its redox potential is *ca.* 100 mV more negative than those of the other phenothiazines. Thus, it may not be surprising that the protocol correctly identifies PT. However, the protocol can also differentiate between MPT and iPrPT, a distinction more challenging to achieve visually, as their redox potentials are much more similar (**Table S3**).

Although the protocol was successful in this case study, additional findings point to limitations and areas for improvement. We note that the protocol did not correctly label voltammograms at faster CV scan rates (200-1000 mV s^{-1}), misidentifying MPT and iPrPT (results illustrated by **Figure S6** in the SI). This misidentification may arise from multiple factors. First, as already noted, the redox potentials of iPrPT and MPT are similar (*ca*. 30 mV separation), potentially frustrating

differentiation using voltammetry and necessitating the use of other techniques that can capitalize on contrasting compound properties. For example, the ^{1}H NMR spectrum of iPrPT exhibits a multiplet at *ca*. 4.25 ppm, which is not present in the analogous spectrum for MPT (see **Figures S10** and **S12** for further details). Further, kinetic limitations manifest themselves to a greater extent at faster CV scan rates *via* increased peak-to-peak separation. If an analyte (e.g., iPrPT in this study) is predicted to undergo an electron transfer with infinitely fast kinetics, as the diffusion rate-limited model in this work assumes, the simulated peak-to-peak separation will be independent of scan rate; however, this separation will increase for analytes predicted to undergo a kinetically rate-limited electron transfer (as is the case with MPT). Consequently, the modeled peak potentials for iPrPT will not change while those for MPT will, creating a greater opportunity for the protocol to confuse iPrPT and MPT. Relatedly, ohmic-induced potential losses distort voltammograms acquired at high scan rates in a similar fashion to that of kinetically rate-limited electron transfers, and as such, ohmic-driven distortions can be misinterpreted as kinetic limitations if appropriate care is not taken. Finally, increased experimental noise and contribution from background charging currents[56,57] (see **Figure S7**) can further convolute the signal and thus analysis. Although data quality appears to impact the ability of the protocol to correctly identify analytes, quantitative metrics of sufficient data quality for accurate labeling were not identified in this study and are expected to be challenging to formulate; they are likely dependent on multiple factors, such as the similarity of the compounds in the library and the type of potentiostat used. Overall, this misidentification demonstrates that the experimental conditions and the validity of the physical models used must be carefully considered.

To increase prediction accuracy, experiments should seek to acquire high-quality (i.e., high signal-noise) data, and the limitations of first-principle models should be actively considered. The

results from this study indicate that cyclic voltammograms should be acquired at a scan rate of *ca.* 100 mV s^{-1} or slower. CSW voltammetry, in turn, did not exhibit analogous limitations, and its threshold waveform inputs are not presently known. Moreover, despite the advantages physical models impart, the experimental conditions of the system being probed as compared to the training set must be evaluated; if there are significant deviations, modified or more detailed physical models may be needed. For example, at faster CV scan rates, double-layer capacitance can appreciably affect the observed current and thus may need to be considered in the physical model,[5] while such effects may not significantly impact CSW voltammetry.[20] More generally, it may not be possible to identify analytes with similar redox potentials using data from a single voltammogram; to this end, complimentary techniques (e.g., UV-Vis spectroscopy, NMR, or additional sensitive voltammetric techniques[40,58]) could be integrated with the Bayesian workflow to increase labeling accuracy in these instances. Such expanded frameworks will be contemplated in due course.

4. Conclusions

In this work, a protocol combining voltammetry experiments and simulations, binary hypothesis testing, and Bayesian inference has been developed to improve the ability to accurately identify analytes in solutions compared to only experiment alone, using experiment combined with simulation, and using black-box (physics-agnostic) machine learning methods. The procedure was outlined and applied to a test case involving five phenothiazines probed with two voltammetry techniques; there, solutions containing PT and MPT were correctly labeled across various techniques (CV and CSW voltammetry). These results demonstrate that a voltammetric labeling protocol can characterize a multi-analyte solution using different techniques, demonstrating a

degree of versatility not yet observed in existing voltammetric identification protocols. Future work will aim to improve the detection accuracy of this methodology by integrating the results of additional techniques in an automated fashion.

Overall, this protocol serves as a first step in extending the limits of electrochemical analysis *via* integration of probabilistic principles with high-quality experimental data (potentially *in situ* or *operando*) and simulations. While the compositions of the solutions examined in this study were known and unchanging, our protocol may ultimately examine more complex and dynamic systems. If validation on these transient systems are promising, this protocol can be used in relevant fields that would benefit from enhanced *in situ* voltammetric labeling; examples include identifying electroactive decay products of degraded analytes during organic redox flow battery operation[59] and labeling potentially complex liquid product mixtures arising from carbon dioxide reduction,[60] both almost in real-time.

5. Acknowledgements

This work was supported as part of the National Science Foundation (NSF) under Award Number 1805566. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. We gratefully acknowledge the MIT Supercloud and Lincoln Laboratory Supercomputing Center for providing HPC resources that have contributed to the research results reported within this paper. We thank Professor Susan Odom, Dr. Aman Kaur, and the Odom Research Group for synthesizing, purifying, analyzing, and shipping phenothiazines. We thank Bertrand Neyhouse, Weiran Gao, Dr. Katharine Greco, Dr. Michael Orella, Kara Rodby, and Dr. Kevin Tenny (all of the Brushett Research Group) for fruitful discussions on data analysis and presentation. We also thank Kindle Williams for her

constructive feedback in revising the manuscript. Finally, we thank Bertrand Neyhouse for assistance in data acquisition.

CRediT authorship contribution statement

Alexis M. Fenton Jr.: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Visualization, Writing – original draft, Writing – review and editing.
Fikile R. Brushett: Conceptualization, Project administration, Supervision, Writing – original draft, Writing – review and editing, Funding acquisition.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

6. Glossary

Latin variables

$\boldsymbol{C}$	Vector of concentrations (mol m^{-3})
$\boldsymbol{C}_{best}$	Vector of best-fit concentrations (mol m^{-3})
$\boldsymbol{C}_{\boldsymbol{best,no\ X}}$	Vector of best-fit concentrations, excluding species X (mol m^{-3})
$C_{i,best}$	Best-fit concentration for species i (mol m^{-3})
C_i	Concentration of species i (mol m^{-3})
$C_{i,bulk}$	Concentration of species i in the bulk (mol m^{-3})
c_i	Dimensionless concentration of species i
D_i	Diffusion coefficient of species i (m^2 s^{-1})
d_O	Ratio of diffusion coefficients $D_O \cdot D_R^{-1}$
E	Applied electrode potential (V vs. reference redox event)
E_0	Formal redox potential for a species of interest (V vs. reference redox event)
F	Faraday constant (96485 C mol^{-1})
$f(\ )$	Continuous probability distribution function, or abbreviation for likelihood function
$\boldsymbol{I}_{\boldsymbol{exp}}$	Vector of experimental currents (A)*
$I_{exp,j}$	j^{th} data point of the experimental current vector (A)
I_i	Current of species i (A)
$\boldsymbol{I}_m$	Vector of modeled currents (A)
$I_{m,j}$	j^{th} data point of the experimental current vector (A)
i	Indexing counter
j	Indexing counter
K_0	Dimensionless heterogeneous rate constant
k_0	Heterogeneous rate constant (m s^{-1})
N	Number of data points in a voltammogram
$P(\ )$	Discrete probability mass function
$P_{background}$	Probability that the current signal arises from non-faradaic processes
$P_f(\ )$	The probability a compound is present in solution when considering background processes
q	Indexing counter
R_G	Universal gas constant (8.314 J mol^{-1} K^{-1})
r_e	Working electrode radius (m)
T	Temperature (K)
t	Time (s)
x	Axial distance from the planar electrode surface (m)

Greek variables

α	Dimensionless charge transfer coefficient
$\boldsymbol{\eta}$	Vector of overpotentials (V)
η	Scalar overpotential (V)
η_j	Overpotential at the j^{th} data point (V)
$\boldsymbol{\theta}$	Generic vector of electrochemical and transport parameters or concentrations (multiple units)
$\boldsymbol{\theta}_i$	Vector of electrochemical and transport parameters for species i (multiple units)
$\hat{\boldsymbol{\theta}}$	Vector of optimal electrochemical and transport parameters or concentrations (multiple units)
ξ	Dimensionless position
$\boldsymbol{\sigma}$	Vector of the standard deviations for the experimental current (A)
σ_j	Standard deviation of the experimental current at the j^{th} data point (A)
τ	Dimensionless time
φ_a	Anodic concentration-normalized difference current (A m^3 mol^{-1})
φ_c	Cathodic concentration-normalized difference current (A m^3 mol^{-1})
$\boldsymbol{\varphi}_i$	Vector of concentration-normalized difference currents for species i (A m^3 mol^{-1})
$\boldsymbol{\varphi}$	Matrix of concentration-normalized difference currents for all species (A m^3 mol^{-1})

*Note that the inclusion of "Δ" before any form of the current or concentration normalized current indicates a difference current or the normalized analog (A or A m^3 mol^{-1}).

Latin symbols

A	Toy analyte used to demonstrate the protocol methodology
a	An anodic (oxidative) process
B	Toy analyte used to demonstrate the protocol methodology
c	A cathodic (reductive) process
H_i	i^{th} hypothesis
M	Total number of hypotheses
O	Oxidized form of a redox couple
O_{obs}	An observation
R	Reduced form of a redox couple
X	A generic analyte
Y	A generic analyte
Z	A generic analyte

7. References

(1) Kwabi, D. G.; Lin, K.; Ji, Y.; Kerr, E. F.; Goulet, M.-A.; De Porcellinis, D.; Tabor, D. P.; Pollack, D. A.; Aspuru-Guzik, A.; Gordon, R. G.; Aziz, M. J. Alkaline Quinone Flow Battery with Long Lifetime at pH 12. *Joule* **2018**, *2* (9), 1894–1906. https://doi.org/10.1016/j.joule.2018.07.005.

(2) O'Dea, J. J.; Osteryoung, J.; Lane, T. Determining Kinetic Parameters from Pulse Voltammetric Data. *J. Phys. Chem.* **1986**, *90* (12), 2761–2764. https://doi.org/10.1021/j100403a040.

(3) Compton, R. G.; Banks, C. E. *Understanding Voltammetry*, 2nd ed.; Imperial College Press: London, 2011.

(4) Frkonja-Kuczin, A.; Alicea-Salas, J. Y.; Arroyo-Currás, N.; Boika, A. Hot-SWV: Square Wave Voltammetry with Hot Microelectrodes. *Anal. Chem.* **2020**, *92*, 8852–8858. https://doi.org/10.1021/acs.analchem.0c00427.

(5) Bard, A. J.; Faulkner, L. R. *Electrochemical Methods: Fundamentals and Applications*, 2nd ed.; John Wiley & Sons. Inc.: New York, 2001.

(6) Quan, M.; Sanchez, D.; Wasylkiw, M. F.; Smith, D. K. Voltammetry of Quinones in Unbuffered Aqueous Solution: Reassessing the Roles of Proton Transfer and Hydrogen Bonding in the Aqueous Electrochemistry of Quinones. *J. Am. Chem. Soc.* **2007**, *129* (8), 12847–12856. https://doi.org/10.1021/ja0743083.

(7) Nicholson, R. S. Theory and Application of Cyclic Voltammetry for Measurement of Electrode Reaction Kinetics. *Anal. Chem.* **1965**, *37* (11), 1351–1355. https://doi.org/10.1021/ac60230a016.

(8) Kowalski, J. A.; Casselman, M. D.; Kaur, A. P.; Milshtein, J. D.; Elliott, C. F.; Modekrutti, S.; Attanayake, N. H.; Zhang, N.; Parkin, S. R.; Risko, C.; Brushett, F. R.; Odom, S. A. A Stable Two-Electron-Donating Phenothiazine for Application in Nonaqueous Redox Flow Batteries. *J. Mater. Chem. A* **2017**, *5*, 24371–24379. https://doi.org/10.1039/C7TA05883G.

(9) Mirčeski, V.; Komorsky-Lovrić, Š.; Lovrić, M. *Square-Wave Voltammetry: Theory and Application*; Springer-Verlag Berlin Heidelberg: Berlin/Heidelberg, 2007.

(10) Kowalski, J. A.; Fenton Jr., A. M.; Neyhouse, B. J.; Brushett, F. R. A Method for Evaluating Soluble Redox Couple Stability Using Microelectrode Voltammetry. *J. Electrochem. Soc.* **2020**, *167*, 160513. https://doi.org/10.1149/1945-7111/abb7e9.

(11) Goulet, M.-A.; Tong, L.; Pollack, D. A.; Tabor, D. P.; Odom, S. A.; Aspuru-Guzik, A.; Kwan, E. E.; Gordon, R. G.; Aziz, M. J. Extending the Lifetime of Organic Flow Batteries via Redox State Management. *J. Am. Chem. Soc.* **2019**, *141* (20), 8014–8019. https://doi.org/10.1021/jacs.8b13295.

(12) Lin, K.; Chen, Q.; Gerhardt, M. R.; Tong, L.; Kim, S. B.; Eisenach, L.; Valle, A. W.; Hardee, D.; Gordon, R. G.; Aziz, M. J.; Marshak, M. P. Alkaline Quinone Flow Battery. *Science* **2015**, *349* (6255), 1529–1532. https://doi.org/10.1126/science.aab3033.

(13) Bond, A. M. Past, Present and Future Contributions of Microelectrodes to Analytical Studies Employing Voltammetric Detection: A Review. *The Analyst* **1994**, *119* (11), 1–21. https://doi.org/10.1039/AN994190001R.

(14) Komorsky-Lovrić, Š. A Simple Method for Detection of Manganese in Marine Sediments. *Croat. Chem. Acta* **1998**, *71* (2), 263–269.

(15) Komorsky-Lovrić, Š.; Vukašinović, N.; Penovski, R. Voltammetric Determination of Microparticles of Some Local Anesthetics and Antithusics Immobilized on the Graphite Electrode. *Electroanalysis* **2003**, *15* (5–6), 544–547. https://doi.org/10.1002/elan.200390067.
(16) Grygar, T.; Kučková, Š.; Hradil, D.; Hradilová, D. Electrochemical Analysis of Natural Solid Organic Dyes and Pigments. *J. Solid State Electrochem.* **2003**, *7* (10), 706–713. https://doi.org/10.1007/s10008-003-0380-1.
(17) Komorsky-Lovrić, Š.; Galić, I.; Penovski, R. Voltammetric Determination of Cocaine Microparticles. *Electroanalysis* **1999**, *11* (2), 120–123. https://doi.org/10.1002/(SICI)1521-4109(199902)11:2<120::AID-ELAN120>3.0.CO;2-R.
(18) Deen, William M. *Analysis of Transport Phenomena*, 2nd ed.; Oxford University Press: New York, 2012.
(19) Compton, R. G.; Laborda, E.; Ward, K. R. *Understanding Voltammetry: Simulation of Electrode Processes*; Imperial College Press: London, 2014.
(20) Helfrick, J. C.; Bottomley, L. A. Cyclic Square Wave Voltammetry of Single and Consecutive Reversible Electron Transfer Reactions. *Anal. Chem.* **2009**, *81* (21), 9041–9047. https://doi.org/10.1021/ac9016874.
(21) O'Dea, J. J.; Osteryoung, J.; Osteryoung, R. A. Theory of Square Wave Voltammetry for Kinetic Systems. *Anal. Chem.* **1981**, *53*, 695–701. https://doi.org/10.1021/ac00227a028.
(22) Oldham, K. B.; Myland, J. C. Modelling Cyclic Voltammetry without Digital Simulation. *Electrochimica Acta* **2011**, *56*, 10612–10625. https://doi.org/10.1016/j.electacta.2011.05.044.
(23) Bond, A. M.; Oldham, K. B.; Zoski, C. G. Theory of Electrochemical Processes at an Inlaid Disc Microelectrode under Steady-State Conditions. *J. Electroanal. Chem.* **1988**, *245*, 71–104. https://doi.org/10.1016/0022-0728(88)80060-3.
(24) *Microelectrodes: Theory and Applications*; Montenegro, M. I., Queirós, M. A., Daschbach, J. L., Eds.; Kluwer Academic Publishers: Dordrecht, 1991.
(25) Narayana, K. A.; Casselman, M. D.; Elliott, C. F.; Ergun, S.; Parkin, S. R.; Risko, C.; Odom, S. A. *N*-Substituted Phenothiazine Derivatives: How the Stability of the Neutral and Radical Cation Forms Affects Overcharge Performance in Lithium-Ion Batteries. *ChemPhysChem* **2015**, *16* (6), 1179–1189. https://doi.org/10.1002/cphc.201402674.
(26) Casselman, M. D.; Kaur, A. P.; Narayana, K. A.; Elliott, C. F.; Risko, C.; Odom, S. A. The Fate of Phenothiazine-Based Redox Shuttles in Lithium-Ion Batteries. *Phys. Chem. Chem. Phys.* **2015**, *17* (10), 6905–6912. https://doi.org/10.1039/c5cp00199d.
(27) Zhao, E. W.; Liu, T.; Jónsson, E.; Lee, J.; Temprano, I.; Jethwa, R. B.; Wang, A.; Smith, H.; Carretero-González, J.; Song, Q.; Grey, C. P. In Situ NMR Metrology Reveals Reaction Mechanisms in Redox Flow Batteries. *Nature* **2020**, *579*, 224–228. https://doi.org/10.1038/s41586-020-2081-7.
(28) Zhao, E. W.; Jónsson, E.; Jethwa, R. B.; Hey, D.; Lyu, D.; Brookfield, A.; Klusener, P. A. A.; Collison, D.; Grey, C. P. Coupled *In Situ* NMR and EPR Studies Reveal the Electron Transfer Rate and Electrolyte Decomposition in Redox Flow Batteries. *J. Am. Chem. Soc.* **2021**, *143* (4), 1885–1895. https://doi.org/10.1021/jacs.0c10650.
(29) Farahani, K. Z.; Benvidi, A.; Rezaeinasab, M.; Abbasi, S.; Abdollahi-Alibeik, M.; Rezaeipoor-Anari, A.; Zarchi, M. A. K.; Abadi, S. S. A. D. M. Potentiality of PARAFAC Approaches for Simultaneous Determination of N-Acetylcysteine and Acetaminophen

Based on the Second-Order Data Obtained from Differential Pulse Voltammetry. *Talanta* **2019**, *192*, 439–447. https://doi.org/10.1016/j.talanta.2018.08.092.
(30) Erickson, J. S.; Shriver-Lake, L. C.; Zabetakis, D.; Stenger, D. A.; Trammell, S. A. A Simple and Inexpensive Electrochemical Assay for the Identification of Nitrogen Containing Explosives in the Field. *Sens. Switz.* **2017**, *17* (8), 1769. https://doi.org/10.3390/s17081769.
(31) Ye, J. J.; Lin, C. H.; Huang, X. J. Analyzing the Anodic Stripping Square Wave Voltammetry of Heavy Metal Ions via Machine Learning: Information beyond a Single Voltammetric Peak. *J. Electroanal. Chem.* **2020**, *872*, 113934. https://doi.org/10.1016/j.jelechem.2020.113934.
(32) Dean, S. N.; Shriver-Lake, L. C.; Stenger, D. A.; Erickson, S.; Golden, J. P.; Trammell, S. A. Machine Learning Techniques for Chemical Identification Using Cyclic Square Wave Voltammetry. *Sens. Switz.* **2019**, *19*, 2392. https://doi.org/10.3390/s19102392.
(33) Shriver-Lake, L. C.; Myers-Ward, R. L.; Dean, S. N.; Erickson, J. S.; Stenger, D. A.; Trammell, S. A. Multilayer Epitaxial Graphene on Silicon Carbide: A Stable Working Electrode for Seawater Samples Spiked with Environmental Contaminants. *Sens. Switz.* **2020**, *20* (14), 4006. https://doi.org/10.3390/s20144006.
(34) de Souza Schaumlöffel, L.; Bolognese Fernandes, P. R.; Sartori Piatnicki, C. M.; Gutterres, M. A Chemometric-Assisted Voltammetric Method for Simultaneous Determination of Four Antioxidants in Biodiesel Samples. *Energy Fuels* **2020**, *34*, 412–418. https://doi.org/10.1021/acs.energyfuels.9b02810.
(35) Bertsekas, D. P.; Tsitsiklis, J. N. *Introduction to Probability*, 2nd ed.; Athena Scientific: Nashua, NH, 2008.
(36) Brandt, R. E.; Kurchin, R. C.; Steinmann, V.; Ceder, G.; Unold, T.; Buonassisi, T. Rapid Photovoltaic Device Characterization through Bayesian Parameter Estimation. *Joule* **2017**, *1*, 843–856. https://doi.org/10.1016/j.joule.2017.10.001.
(37) Wang, Y.; Kaur, A. P.; Attanayake, N. H.; Yu, Z.; Suduwella, T. M.; Cheng, L.; Odom, S. A.; Ewoldt, R. H. Viscous Flow Properties and Hydrodynamic Diameter of Phenothiazine-Based Redox-Active Molecules in Different Supporting Salt Environments. *Phys. Fluids* **2020**, *32* (8), 083108. https://doi.org/10.1063/5.0010168.
(38) Winslow, S. W.; Shcherbakov-Wu, W.; Liu, Y.; Tisdale, W. A.; Swan, J. W. Characterization of Colloidal Nanocrystal Surface Structure Using Small Angle Neutron Scattering and Efficient Bayesian Parameter Estimation. *J. Chem. Phys.* **2019**, *150* (24), 224702. https://doi.org/10.1063/1.5108904.
(39) 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*, 2041–2050. https://doi.org/10.1007/s10008-020-04639-6.
(40) Gavaghan, D. J.; Cooper, J.; Daly, A. C.; Gill, C.; Gillow, K.; Robinson, M.; Simonov, A. N.; Zhang, J.; Bond, A. M. Use of Bayesian Inference for Parameter Recovery in DC and AC Votlammetry. *ChemElectroChem* **2018**, *5* (6), 917–935. https://doi.org/10.1002/celc.201700678.
(41) Li, J.; Kennedy, G. F.; Gundry, L.; Bond, A. M.; Zhang, J. Application of Bayesian Inference in Fourier-Transformed Alternating Current Voltammetry for Electrode Kinetic Mechanism Distinction. *Anal. Chem.* **2019**, *91*, 5303–5309. https://doi.org/10.1021/acs.analchem.9b00129.

(42) Northrop, P. W. C.; Cole, J. V. A Pulse Voltammetry Analysis Toolkit for Battery and Fuel Cell Material. *ECS Trans.* **2018**, *85* (5), 23–42. https://doi.org/10.1149/08505.0023ecst.
(43) Limaye, A. M.; Zeng, J. S.; Willard, A. P.; Manthiram, K. Bayesian Data Analysis Reveals No Preference for Cardinal Tafel Slopes in CO_2 Reduction Electrocatalysis. *Nat. Commun.* **2021**, *12*, 703. https://doi.org/10.1038/s41467-021-20924-y.
(44) Gundry, L.; Guo, S.-X.; Kennedy, G. F.; Keith, J.; Robinson, M.; Gavaghan, D. J.; Bond, A. M.; Zhang, J. Recent Advances for Future Perspectives for Automated Parameterisation, Bayesian Inference and Machine Learning in Voltammetry. *Chem. Commun.* **2021**, *57* (15), 1855–1870. https://doi.org/10.1039/d0cc07549c.
(45) Gagné, R. R.; Koval, C. A.; Lisensky, G. C. Ferrocene as an Internal Standard for Electrochemical Measurements. *Inorg. Chem.* **1980**, *19*, 2855–2857. https://doi.org/10.1021/ic50211a080.
(46) EC-Lab® Software: Techniques and Applications - Version 10.38. BioLogic Science Instruments August 2014.
(47) Kim, E.; Huang, K.; Saunders, A.; McCallum, A.; Ceder, G.; Olivetti, E. Materials Synthesis Insights from Scientific Literature via Text Extraction and Machine Learning. *Chem. Mater.* **2017**, *29*, 9436–9444. https://doi.org/10.1021/acs.chemmater.7b03500.
(48) Chin, K. Y.; Prasad, S.; O'Dea, J. J.; Osteryoung, J. Mathematical Enhancement of the Performance of Voltammetric Sensors. *Anal. Chim. Acta* **1992**, *264* (2), 197–204. https://doi.org/10.1016/0003-2670(92)87006-7.
(49) Pellitero, M. A.; Shaver, A.; Arroyo-Currás, N. Critical Review—Approaches for the Electrochemical Interrogation of DNA-Based Sensors: A Critical Review. *J. Electrochem. Soc.* **2020**, *167*, 037529. https://doi.org/10.1149/2.0292003JES.
(50) Helfrick, Jr., J. C.; Mann, M. A.; Bottomley, L. A. Diagnostic Criteria for the Characterization of Electrode Reactions with Chemical Reactions Following Electron Transfer by Cyclic Square Wave Voltammetry. *Electrochimica Acta* **2016**, *205*, 20–28. https://doi.org/10.1016/j.electacta.2016.04.006.
(51) Reuther, A.; Kepner, J.; Byun, C.; Samsi, S.; Arcand, W.; Bestor, D.; Bergeron, B.; Gadepally, V.; Houle, M.; Hubbel, M.; Jones, M.; Klein, A.; Milechin, L.; Mullen, J.; Prout, A.; Rosa, A.; Yee, C.; Michaleas, P. Interactive Supercomputing on 40,000 Cores for Machine Learning and Data Analysis; IEEE: Waltham, MA, 2018; pp 1–6. https://doi.org/10.1109/HPEC.2018.8547629.
(52) Gunasekara, I.; Mukerjee, S.; Plichta, E. J.; Hendrickson, M. A.; Abraham, K. M. Microelectrode Diagnostics of Lithium-Air Batteries. *J. Electrochem. Soc.* **2014**, *161* (3), A381–A392. https://doi.org/10.1149/2.073403jes.
(53) Buhrmester, C.; Moshurchak, L.; Wang, R. L.; Dahn, J. R. Phenothiazine Molecules: Possible Redox Shuttle Additives for Chemical Overcharge and Overdischarge Protection for Lithium-Ion Batteries. *J. Electrochem. Soc.* **2006**, *153* (2), A288–A294. https://doi.org/10.1149/1.2140615.
(54) Attanayake, N. H.; Kowalski, J. A.; Greco, K. V.; Casselman, M. D.; Milshtein, J. D.; Chapman, S. J.; Brushett, F. R.; Odom, S. A. Tailoring Two-Electron-Donating Phenothiazines To Enable High-Concentration Redox Electrolytes for Use in Nonaqueous Redox Flow Batteries. *Chem. Mater.* **2019**, *31* (12), 4353–4363. https://doi.org/10.1021/acs.chemmater.8b04770.

(55) Odom, S. A.; Ergun, S.; Poudel, P. P.; Parkin, S. R. A Fast, Inexpensive Method for Predicting Overcharge Performance in Lithium-Ion Batteries. *Energy Environ. Sci.* **2014**, *7*, 760–767. https://doi.org/10.1039/C3EE42305K.
(56) Roberts, J. G.; Sombers, L. A. Fast-Scan Cyclic Voltammetry: Chemical Sensing in the Brain and Beyond. *Anal. Chem.* **2018**, *90*, 490–504. https://doi.org/10.1021/acs.analchem.7b04732.
(57) Venton, B. J.; Cao, Q. Fundamentals of Fast-Scan Cyclic Voltammetry for Dopamine Detection. *Analyst* **2020**, *145*, 1158–1168. https://doi.org/10.1039/c9an01586h.
(58) Robinson, M.; Ounnunkad, K.; Zhang, J.; Gavaghan, D.; Bond, A. Integration of Heuristic and Automated Parametrization of Three Unresolved Two-Electron Surface-Confined Polyoxometalate Reduction Processes by AC Voltammetry. *ChemElectroChem* **2018**, *5* (23), 3771–3785. https://doi.org/10.1002/celc.201800950.
(59) Kwabi, D. G.; Ji, Y.; Aziz, M. J. Electrolyte Lifetime in Aqueous Organic Redox Flow Batteries: A Critical Review. *Chem. Rev.* **2020**, *120* (14), 6467–6489. https://doi.org/10.1021/acs.chemrev.9b00599.
(60) Leonard, M. E.; Orella, M. J.; Aiello, N.; Román-Leshkov, Y.; Forner-Cuenca, A.; Brushett, F. R. Flooded by Success: On the Role of Electrode Wettability in CO_2 Electrolyzers That Generate Liquid Products. *J. Electrochem. Soc.* **2020**, *167*, 124521. https://doi.org/10.1149/1945-7111/abaa1a.