Article:
Lloyd-Laney, Henry O, Yates, Nicholas D J, Robinson, Martin J et al. (7 more authors) (2021) Using Purely Sinusoidal Voltammetry for Rapid Inference of Surface-Confined Electrochemical Reaction Parameters. Analytical Chemistry. pp. 2062-2071. ISSN 0003-2700

https://doi.org/10.1021/acs.analchem.0c03774

Using Purely Sinusoidal Voltammetry for Rapid Inference of Surface-Confined Electrochemical Reaction Parameters

Henry O. Lloyd-Laney,[†] Nicholas D. J. Yates,[‡] Martin J. Robinson,[†] Alice R. Hewson,[‡] Jack D. Firth,[‡] Darrell M. Elton,[¶] Jie Zhang,[§] Alan M. Bond,[§] Alison Parkin,[*,‡] and David J. Gavaghan[*,†]

[†]*Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD United Kingdom*

[‡]*Department of Chemistry, University of York, Heslington, York, YO10 5DD, United Kingdom*

[¶]*School of Engineering and Mathematical Sciences, La Trobe University, Bundoora, Victoria 3086, Australia*

[§]*School of Chemistry and the ARC Centre of Excellence for Electromaterials Science, Monash University, Clayton, Vic., 3800 Australia*

E-mail: alison.parkin@york.ac.uk; david.gavaghan@dtc.ox.ac.uk

Abstract

Alternating current (AC) voltammetric techniques are experimentally powerful as they enable Faradaic current to be isolated from non-Faradaic contributions. To find the best global fit between experimental voltammetric data and simulations based on reaction models requires searching a substantial parameter space at high resolution. In this paper we estimate parameters from purely sinusoidal voltammetry (PSV) experiments, investigating the redox reactions of a surface-confined ferrocene derivative. The advantage of PSV is that a complete experiment can be simulated relatively rapidly, compared to other AC voltammetric techniques. In one example involving thermodynamic dispersion, a PSV parameter inference effort requiring 7,500,000 simulations was completed in 7 hours whereas the same process for our previously used technique, ramped Fourier transform AC voltammetry (ramped FTACV) would have taken four days. Using both synthetic and experimental data with a surface confined diazonium substituted ferrocene derivative, it is shown that the PSV technique can be used to recover the key chemical and physical parameters. By applying techniques from Bayesian inference and Markov chain Monte-Carlo methods, the confidence, distribution and degree of correlation of the recovered parameters was visualised and quantified.

Introduction

Electron transfer reactions, where the oxidation of one species is accompanied by the reduction of another, are essential to life.[1] These so-called redox reactions also underpin battery technology, fuel cell chemistry, synthetic electrochemistry and many other modern technologies.[2,3] Voltammetry, a key technique in the study of redox chemistry, requires the measurement and interpretation of current-potential-time ($\tilde{I}$-$\tilde{E}$-$\tilde{t}$) data. Quantitative analysis usually involves inferring chemical reaction parameters using

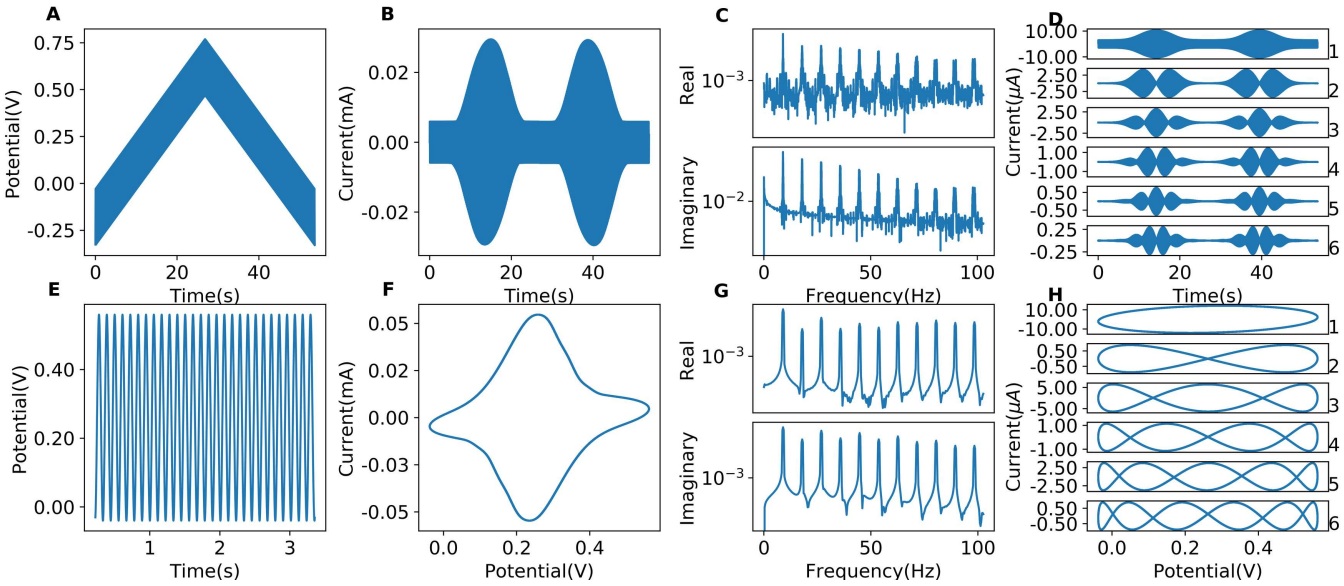

Figure 1: Comparison of the output of (top) ramped and (bottom) purely sinusoidal voltammograms simulated using identical model parameters (see table S1). (A and E) Time dependence of the applied waveform. (B and F) Total current response to inputs in A and E, respectively. (C and G) Absolute power spectra derived by Fourier transformation of B and D, respectively. (D and H) Resolved 1-6th harmonics created by band selection on the harmonics in C and G followed by inverse Fourier transform.

models that are proposed to mimic the electrochemical response.[4] Typically, the mathematics used to simulate the voltammetric experiment consist of a system of non-linear partial differential equations for solution-phase electrochemistry,[5] where the analyte is dissolved in the electrolyte, or non-linear ordinary differential equations for "film" electrochemistry,[6] where the redox-active analyte is confined to the surface of the working electrode. Embedded within these models are the chemical reaction parameters that describe the Faradaic current that arises from electron transfer process(es), such as the rate constants of all reactions taking place, the corresponding equilibrium potentials and transfer coefficients.[7] Models must also contain parameters describing the impact of uncompensated resistance as well as non-Faradaic background current that arises from electrode-solution capacitive charging.[7]

Commonly, these mathematical models cannot be solved analytically, and hence computationally supported simulation approaches are necessary.[5] The question then arises as to how we best use the simulations to estimate, from experimental data, the parameters governing the electrochemical system being investigated. This exercise is known as the inverse problem, where a large number of simulations are made with different combinations of parameters, with the purpose of finding the minimum distance between simulation and experiment. This can be a computationally intensive exercise where the computing time consumed will be voltammetric technique dependent; the computing time for the actual optimisation algorithm is largely independent of technique. In the most complex problem we have addressed by data optimization, over 20 parameters were inferred from a single alternating current (AC) voltammetric data set.[8] However, a significant limitation with the data optimization approach cited is that only point estimates of the parameter values are generated, whereas detailed understanding also requires knowledge of the accuracy of each inferred parameter value, although it should be noted that the Bayesian framework we use in this current study is one method amongst many of obtaining this knowledge,[9] and a more complete review of these methods can be found in our previous work.[10]

Even though more effort to address the inverse problem by employing data optimization would be invaluable, it is also apparent that even more sophisticated and powerful parameter inference methods should probably be introduced.[11] For example, recently,[10,12] we examined the use of Bayesian inference as a tool for addressing the inverse problem, an approach which is very commonly used in other fields of science.[13,14] We demonstrated how a Bayesian approach enables the use of computational methods such as Markov Chain Monte Carlo (MCMC) to provide an estimate of both the optimal parameter values and the likely spread of those parameters about the optimum. We illustrated the power of this approach by demonstrating that the technique of AC voltammetry[15] is more sensitive to differences in chemical reaction parameters compared to the simpler but more widely used DC technique. In the version of AC voltammetry used in the referenced work, the applied waveform was the sum of a linear ramp, as in a DC voltammetric experiment, and a sine wave having a frequency f (see figure 1). Fourier transforming the total current response allows the signal to be resolved into a series of harmonics detected at $f, 2f, 3f, \ldots, nf$, as well as the aperiodic DC component (not shown). This version of AC voltammetry is known as Fourier Transformed AC Voltammetry or FTACV. Access to the mechanistically sensitive higher order AC harmonics provides Faradaic responses devoid of background charging current.[15–17] To distinguish this technique from other forms of voltammetry based on sinusoidal waveforms, herein AC voltammetry with an underlying DC ramped potential is referred to as the "ramped" FTACV method.

The electrode process considered in our initial Bayesian inference study[10] was the quasi-reversible one-electron reduction of $[Fe(CN)_6]^{3-}$ to $[Fe(CN)_6]^{4-}$ in aqueous KCl electrolyte media. In this example, because of the 1 mM concentration of $[Fe(CN)_6]^{3-}$ employed, the total DC plus AC current output is almost purely Faradaic in nature and relatively insensitive to the impact of experimental noise, double layer capacitance and uncompensated resistance. This facilitates very good parameter identifiability for the five system parameters of interest, allowing the inverse problem to be solved straightforwardly and robustly with Bayesian inference. These parameters were the reversible potential, kinetic rate constant, linear double-layer capacitance, uncompensated resistance and symmetry factor (E^0, k_0, C_{dl}, R_u and α respectively). However, there is also widespread interest in voltammetric investigations of surface confined redox-active biological macromolecules such as electron-transfer proteins and oxidoreductase enzymes.[6,18,19] For these systems, typically, the electrode kinetics are fast, thermodynamic and kinetic dispersion is present and surface coverages are low. The high computational costs associated with ramped FTACV simulations make it difficult to apply the Bayesian inference approach to parameterize the higher order harmonic data sets on a reasonable timescale ($\leq$1 day). This high computational cost is exacerbated because relatively high frequencies (>100 Hz) are required to interrogate fast electrode kinetics which requires very fine discretization to solve the underlying differential equations with sufficient accuracy. When thermodynamic and kinetic dispersion also must be modelled,[20] the time taken for a single fitting run can be well over 24 hours.[6] Nevertheless, given sufficient computing power and time, it is of course still possible to infer parameters by ramped FTACV methods.[6,21]

In this paper, specifically aimed at lowering computational requirements in AC voltammetric data analysis of complex problems, the underlying DC ramp present in previous work has been removed. Thus, the potential input is solely based on a large amplitude sine-wave used in what we term as ramp-free "purely sinusoidal voltammetry" or PSV (see figure 1). The PSV method has many appealing properties; the simulation speed is independent of frequency, allowing for interrogation of fast kinetics, the current response is non-linear as in ramped-FTACV, but an entire experiment can be simulated in significantly less time, and the

potential input is entirely continuous. Simulation of 30 oscillations of a PSV experiment as in this study took *0.01s* on a laptop (8GB RAM and an Intel i5-8350U CPU @ 1.70GHz), which represents a 10-15 fold reduction relative to that required for ramped FTACV.

The original form of the PSV method on which the present work is based was invented a long time ago. As noted by Bard and Zoski, in the 1940s fast linear potential sweeps were applied at the dropping mercury electrode, known as oscillographic polarography, as reviewed in.[22] Other researchers have investigated PSV from a theoretical standpoint for a reversible process,[23] and as a means of discrimination against background current relative to the linear sweep DC method.[24] The PSV method has also been used for the sensing of small molecules such as nucleotides and neurotransmitters.[25,26] In a modelling context, large amplitude PSV has been used to infer electrochemical parameters from a voltammetric experiment[27] using background subtraction. Thus, in terms of mechanistic studies, the PSV technique has been predominantly only of theoretical interest with respect to reversible processes, rather than subjected to widespread studies of electrode kinetics, as with the ramped FTACV method. PSV is also related to electrochemical impedance spectroscopy (EIS) which uses multiple small-amplitude phase-randomised sinusoids. In the future we plan to extend our modelling approach to consider parameter estimation in EIS. This will paralell previous work that explored the link between ramped-FTACV and EIS.[15,28]

To demonstrate the significant improvement in time taken to infer parameters from PSV experimental data when solving the inverse problem in a Bayesian framework, we have estimated parameters describing the one-electron oxidation of a diazonium substituted ferrocene derivative surface confined to a glassy carbon (GC) electrode (details provided in Experimental Section) which is shown to be complicated by non-linear capacitance and thermodynamic dispersion. This model system therefore provides a sufficiently challenging problem to illustrate the computation efficiency advantage in the PSV approach.

Mathematical Model

The quasi-reversible one-electron oxidation process of interest in this study is modelled as in our previous work,[6,21] where the total experimental current is the solution to a differential algebraic equation, although it should be noted that to solve it numerically we rearrange it into an ordinary differential equation. A description of the mathematical approach can be found in the SI. The key Faradaic reaction parameters are E^0, k_0 and α, representing the reversible potential, rate constant and symmetry factor respectively, as defined by the Butler-Volmer equation. In order to solve for the total current numerically, we discretize the time interval $[0, T]$ into N_t time steps and replace the time derivatives by finite difference approximations. The validity of the numerical approach for the purely sinusoidal case was tested by comparing the solution to an analytical approximation for a reversible surface-confined process responding to a large-amplitude sinusoidal potential input, as provided in the SI in figure S15.[23]

Dispersion

The mathematical model we cite above assumes that the voltammetry of all electroactive molecules can be modelled using the same thermodynamic and electrode kinetic parameters. However, this may not be a valid assumption,[20] especially when the molecules are grafted to an inhomogeneous roughly abraded glassy carbon surface,[20] as in our experiments. In this situation, variability in the orientation of molecules on the surface, or in the underlying surface chemistry may result in a distribution of the parameters.[20] We therefore include a measure of dispersion in the calculation of I, where the selected parameters are modelled as a distribution instead of as a point estimate. Discussion of how this is achieved mathematically is provided in the SI, and in figure S1. According to experimental observations,[20,29] when dispersion

is known to be present for the parameters E^0 and k_0, the distributions are normal and lognormal respectively. There is no information on the distribution of the α parameter, but as it is also related to the energetics of the reaction (specifically the energy of the transition state), the effect of introducing a distribution in this parameter in a similar fashion to that of the thermodynamic parameter E^0 also is explored. Each dispersed parameter requires N simulations to approximate the distribution, and consequently the time taken to simulate the forward problem for n dispersed parameters is N^n, which was a key motivation for reducing simulation time for parameter estimation for voltammetric experiments.

Methods

Electrochemistry experiments

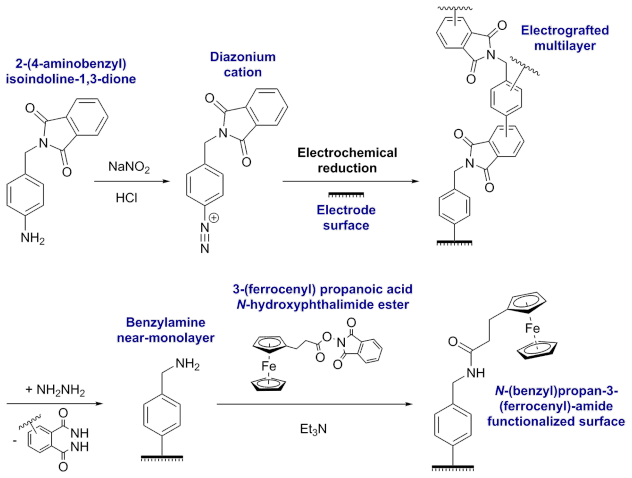

Figure 2: Cartoon representing the diazonium functionalization procedure used to generate a ferrocene-coated glassy carbon electrode surface.

A 3.0 mm diameter stationary glassy carbon disk working electrode (eDAQ) was used in all voltammetric experiments. As summarized in figure 2 and detailed in the SI, in situ diazonium cation generation and electrografting was carried out using 2-(4-aminobenzyl)isoindoline-1,3-dione to generate an electrografted multilayer on the glassy carbon surface. Following a published procedure,[30,31] hydrazine (NH_2NH_2) deprotection was used to reveal the benzylamine functionalities grafted onto the glassy carbon surface. Finally, incubation of the amine-modified electrode in a solution of 3-(ferrocenyl)propanoic acid N-hydroxyphthalimide ester generates a glassy carbon electrde surface covered with immobilized ferrocene moieties, as previously shown in the literature (SI and figure 2).[30] Voltammetric experiments probing the surface-ferrocene electron transfer chemistry were then conducted as detailed in the SI. Potentials in figures are reported as V or mV vs the reference, which is also detailed in the SI.

Synthetic data

In order to assess the ability of our Bayesian inference methodology to recover accurate parameter values for a surface confined process, we initially inferred parameters from synthetic data with Gaussian noise added at each time point. As with our Bayesian inference analysis with solution soluble Ox and Red species,[10] we assume that the added noise has zero mean and a standard deviation that is 0.5-2% of the maximum current value, in line with the observed noise level from experimental work. We denote the current with added noise as $i_{sim}(\tau_j)$, with

$$I_{sim}(t_j) = I(t_j) + \xi_j \tag{1}$$

where $I(t_j)$ is the simulated solution to the forward problem and ξ_j is the normally distributed added noise. This noisy data is then used to estimate or infer the values of the governing parameters that were used to generate the noise-free current trace i.e. to infer the values of $(E^0, k_0, C_{dl}, R_u, \sigma)$, where σ is the inferred standard deviation of the noise. This inference is achieved by setting up the inverse problem as a least squares minimization problem which is optimized to obtain initial point estimates of the parameters using the CMAES algorithm, with estimates of the posterior distributions of those parameters subsequently being obtained by use of a Markov Chain Monte Carlo algorithm. This approach is outlined below, and is described in detail in previous work.[21,32] The

results of these processes were used to quantify and compare the ability of the purely sinusoidal and ramped FTACV methods to determine the true parameter values.

Minimization

In order to infer point values for the parameters defined above we use the covariance matrix adaptation evolution strategy (CMAES) algorithm[33] to minimize an objective function, $\mathcal{L}_t$ by making steps within a bounded parameter space, where the bounds are chemically informed. The algorithm is a derivative-free method to minimize the distance between the data, I_{data} and the noisy simulated time-series $I_{sim}(\tau)$, where N_t is the number of points, i.e. we minimize

$$\mathscr{L}_t = \sqrt{\sum_{j=1}^{N_t} (I_{data}(\tau_j) - I_{sim}(\tau_j))^2}. \qquad (2)$$

Bayesian inference

Once we obtain point estimates for the parameters using the CMAES algorithm, we used Markov-chain Monte-Carlo (MCMC) methods to more fully investigate the solutions we have obtained.[10] The objective function in equation 2 also can be thought of as finding the maximum likelihood estimate, and so we modify it appropriately for use as a likelihood function. We use the adaptive covariance MCMC method, using the PINTS GitHub repository[34] (which, for convenience, we refer to hereafter as MCMC), which attempts to explore the parameter space in a direction that increases the likelihood. This takes the form of repeatedly sampling from parameter space for a set number of iterations ("samples") using 3 independent processes ("chains"). The validity of an inferred distribution is implied by convergence (i.e. multiple chains exploring the same area of parameter space after an initial period). Once the algorithm has reached the pre-defined quantity of samples, we discard a set number of earlier samples that explored lower-likelihood areas of parameter space as "burn-in".[35] The remaining samples of each chain are combined to form the sampling distribution for each parameter — from these distributions we can report the mean, standard deviation and degree of correlation between the various parameters. A more complete description of the algorithm is provided in the GitHub repository associated with this paper.

Results and Discussion

Modelling the parameter impact on PSV

In order to show the visual impact of the effect of the key model parameters, the PSV (current-potential) plots shown in figure 3 were generated, which represent the 2-30th potential oscillation of the sinusoidal input, discarding the first transient oscillation. In each panel, a designated parameter is varied while the others are held constant at values provided in table S2 that were chosen to clearly indicate the effect of each parameter. The symbols μ and σ represent the mean and standard deviation of a normally-dispersed parameter. Actual experimental data for the oxidation of the surface confined ferrocene derivative (which can be found in figure 5) have a lower ratio of Faradaic to capacitance current.

The mean value for the reversible potential, $E^0\mu$, alters the potential at which the Faradaic process is observed and increasing the standard deviation of the reversible potential, $E^0\sigma$, results in peak broadening. The effect of the electrode kinetic term k_0 depends on which kinetic regime (reversible, irreversible or quasi-reversible) the process is in, but in the quasi-reversible regime, the value k_0 significantly affects the Faradaic current peak width and height. The Faradaic response is relatively insensitive to large changes in the value of the uncompensated resistance R_u but it should be noted that the Ohmic IR_u drop term can be significant, and exert similar changes in voltammograms to the k_0, meaning that care is needed to distinguish the impact of each parameter. Capacitance background current is observed as

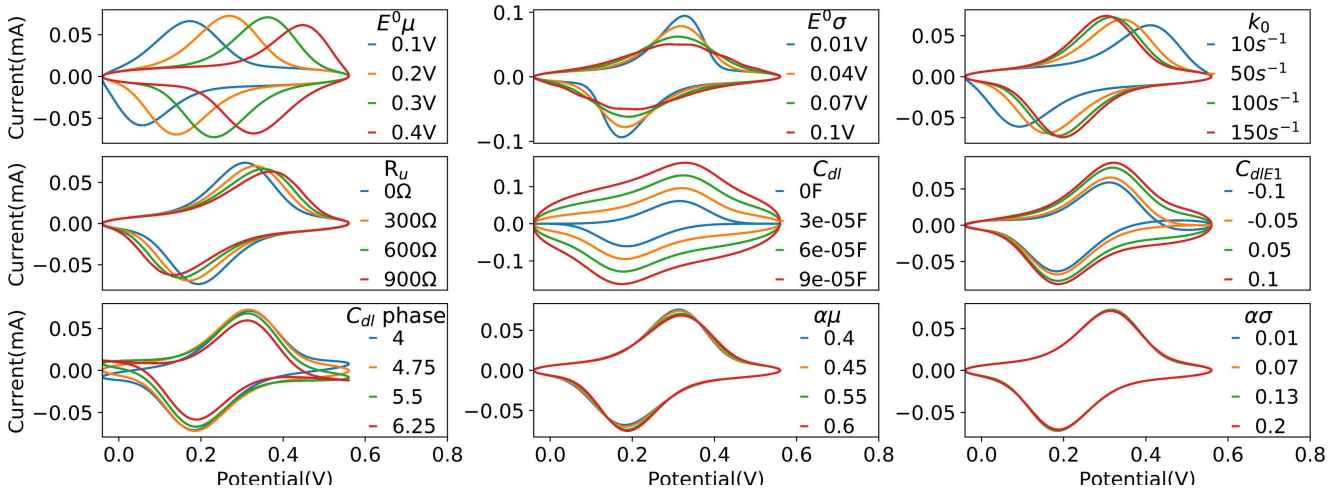

Figure 3: Sensitivity of the simulated PSV response to 9 model parameters. These are, from left to right, the reversible potential mean, $E^0\mu$, and standard deviation $E^0\sigma$, the rate constant k_0, the uncompensated resistance R_u, capacitance parameters, C_{dl}, C_{dlE1} and their associated phase, and the mean and standard deviation of the symmetry factor $\alpha\mu$ and $\alpha\sigma$. The parameter of interest is designated in each panel while the other model parameters are held constant at values reported in table S2.

an ellipse with a current magnitude determined by the value of the C_{dl} parameter — this ellipse can entirely obscure the Faradaic process. Inclusion of higher order capacitance terms such as C_{dlE1} modifies the shape of the ellipse. With the inclusion of an explicit capacitance phase, the possibility of using phase selective detection to resolve the non Faradaic and Faradaic current components is introduced. Finally, while variation in the $\alpha\mu$ parameter modifies the ratio between the peak heights of the oxidation and reduction components slightly, the impact on current magnitude is not very sensitive to changes in $\alpha\sigma$ with the values of k_0 considered. The minimal influence of α also applies to the PSV harmonics within the k_0 range examined, as shown in figure S18. Changing $\alpha\mu$ results in greater changes in the even than odd harmonics in the Fourier spectrum (such as in figure 1(C)). The magnitudes of even harmonics also are much smaller than for the odd harmonics, and therefore contribute less to the overall "shape" of the time series. The effect on the real and imaginary portions of the Fourier spectrum of the other parameters varied systematically in figure 3 is displayed in figures S16 and S17.

Inferring parameters from data containing synthetic noise

To demonstrate that the CMAES algorithm is able to successfully infer parameter distributions from PSV data, a simulated data set, to which noise had been added in order to mimic experimental data, was analysed. Figure 4 shows the results of the parameter inference data generated using both the PSV and ramped FTACV methods, with 0.5% noise added. This translates to a root mean squared error value relative to the original simulation of 4.5 μA. The error observed in experimental data vs the fitted simulation, as inferred by the CMAES algorithm in a later section, was 0.72-1.00 μA, which validates our choice of noise value. The histograms are generated from three chains run independently for 10000 samples, and had 5000 samples discarded as burn-in as described above in the methods section. The ratio of the standard deviations of the inferred parameter distributions for the two techniques for varying amounts of added noise can be found in table S3. The results clearly indicate that it is possible to recover the correct answer from noisy synthetically generated data using both methods, although PSV provides less precise

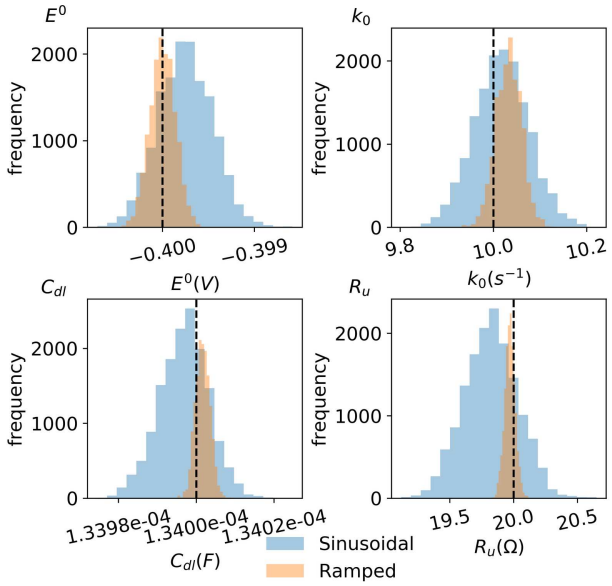

Figure 4: Comparison of inferred parameter distributions (indicated by the symbol) for the purely sinusoidal and ramped experiments, performed on synthetic noisy data derived from a simulated experiment with independent and identically distributed random variables (I.I.D). The simulated data was generated using Normally distributed noise values added to simulated data, where the distribution had 0 mean and a standard deviation that was 0.5% of the maximum current amplitude, and α was set at 0.5. Histograms were created by running the adaptive Metropolis-Hastings algorithm for 10000 samples, where the initial point was the correct parameter value (shown by the black dotted line). An uninformative prior was used, with upper lower bounds of ±50% of the true value. Three independent Markov chains were run, the first 5000 samples discarded as burn-in, and the rest of the samples used to plot the histogram shown.

estimates (as observed by the increased width of distribution) than ramped FTACV. This is quantified in table S3; for the 0.5% noise case, the standard deviations of the parameter distributions inferred from synthetic PSV data is 2-5 times larger than for synthetic ramped data. However, the trade off is the significant decrease in computing time required in the PSV simulations which requires about tenfold fewer points than the ramped FTACV ones; the computation time is approximately reduced by a factor of 12. The time taken for a 100000 sample MCMC run with two dispersed parameters, requiring 7,500,000 solutions to the forward problem took 7 hours; the equivalent run for ramped FTACV would have taken approximately 84 hours, or 3 and a half days.

Inferring parameters from experimental data

A glassy carbon electrode modified with the ferrocene diazonium derivative was subjected to ten PSV experiments followed by ten separate ramped FTACV ones. The PSV experiments used a sinusoidal waveform with a frequency 8.94 Hz and an amplitude 300 mV over a potential range of -40 to +560 mV vs Ref for a time period of 26.8 s, equivalent to 240 sine-wave oscillations. For all PSV inference attempts, we compared simulations to the total current response. For ramped FTACV, we compared simulations to a portion of the Fourier transform of the total current, corresponding to harmonics 2-6. When comparing experimental and simulated PSV currents using equation 2, the initial transient current (corresponding to the first oscillation of the potential input) was discarded, and for PSV data-simulation comparisons, the current was truncated to the times before the 30th oscillation. These manipulations represent a compromise between reducing the number of points for simulation speed, and to ensure that the results of parameter inference attempts were consistent — we found that truncating the data to under 20 oscillations resulted in a loss of convergence between Markov chains when analysing the data using MCMC. By comparison, the AC component of the ramped experiment requires approximately 475 oscillations for complete simulation.

An example result from the parameter inference analysis process is shown in figure 5 for the tenth non- ramped PSV experiment, selected from the ten total repeats, performed consecutively. Because of the close similarity of the simulated and experimental currents, the experimental (orange) line has been made partially transparent, and as such regions of very

Table 1: Inferred CMAES parameters for PSV experiments 1-10 and the first ramped experiment. † The phase and the capacitance phase were both held at 0 during the fitting process. *The error reported for the ramped experiment is the mean error across the fitted harmonics (i.e. 2-6), and so is not directly comparable to the error for the PSV experiments, which is the error for the total timeseries.

Parameter	Symbol	PSV 1	PSV 2	PSV 3	PSV 4	PSV 5	PSV 6	PSV 7	PSV 8	PSV 9	PSV 10	Ramped 1
Midpoint potential mean	$E^0\mu$ (V vs. Ref)	0.247	0.244	0.244	0.244	0.242	0.241	0.241	0.241	0.24	0.241	0.229
Midpoint potential standard deviation	$E^0\sigma$ (V)	0.028	0.047	0.045	0.046	0.047	0.048	0.048	0.048	0.048	0.048	0.046
Rate constant	$k_0(s^{-1})$	104.61	144.066	142.214	142.928	140.303	139.955	141.737	136.411	133.439	136.212	123.33
Uncompensated resistance	Ru (Ω)	732.021	517.171	539.316	540.249	516.33	495.788	491.333	493.475	488.66	499.132	873.541
Linear double-layer capacitance	C_{dl} (F)	7.689E-5	7.681E-5	7.636E-5	7.609E-5	7.609E-5	7.613E-5	7.602E-5	7.582E-5	7.571E-5	7.537E-5	3.341E-5
1^{st} order C_{dl}	C_{dlE1}	3.207E-3	3.294E-3	3.137E-3	2.976E-3	2.738E-3	2.541E-3	2.475E-3	2.391E-3	2.221E-3	2.277E-3	0.058
2^{nd} order C_{dl}	C_{dlE2}	-3.993E-4	-4.402E-4	-4.265E-4	-4.205E-4	-4.219E-4	-4.226E-4	-4.228E-4	-4.195E-4	-4.153E-4	-4.106E-4	-2.122E-3
Surface coverage	$\Gamma(molcm^{-2})$	7.637E-11	7.335E-11	7.330E-11	7.308E-11	7.288E-11	7.271E-11	7.241E-11	7.220E-11	7.205E-11	7.166E-11	7.178E-11
Potential frequency	ω (Hz)	8.941	8.941	8.941	8.941	8.941	8.941	8.94	8.941	8.941	8.941	8.885
C_{dl} phase	C_{dl} phase (rads)	4.392	4.335	4.342	4.342	4.335	4.329	4.328	4.328	4.326	4.328	0.00†
Phase	Phase (rads)	5.082	4.956	4.969	4.968	4.958	4.95	4.947	4.95	4.949	4.951	0.0 †
Symmetry factor mean	$\alpha\mu$	0.588	0.614	0.61	0.609	0.604	0.603	0.603	0.599	0.596	0.599	0.438
Symmetry factor standard deviation	$\alpha\sigma$	0.163	0.188	0.188	0.189	0.19	0.192	0.192	0.191	0.191	0.191	0.155
Error	RMSE (μA)	0.997	0.822	0.837	0.828	0.782	0.753	0.745	0.742	0.722	0.747	0.337*

good agreement are brown. Parameters were obtained by running the CMAES algorithm 10 times for each data set, with chemically informed upper and lower bounds for each parameter, which are reported in the SI in table S4, and noting the best scoring parameter vectors. The initial parameter values for each run of the CMAES algorithm were obtained by randomly sampling from within the boundaries for each value. We believe that these values, presented in table 1 represent the global minimum (at least as defined by the parameter boundaries), as the 10 sets of random initializations will have explored a large portion of parameter space. On average, the solution vectors presented in table 1 were returned by the CMAES algorithm 4-6 times. Standard deviations, as determined by MCMC, can be found in table S7 in the SI. From figure 5 the simulation is seen to provide a good fit to the data in both the time and frequency domains, except in the even harmonics, especially the fourth harmonic, where the simulation overestimates the magnitude relative to the data. We believe that this is because the even harmonics contribute very little to the total signal (note the magnitude of the peaks in the Fourier spectrum, and the size of the current response in the harmonic plots), and are therefore difficult to fit when using the objective function in equation 2. It should be noted that the inaccuracies in the fit (such as overstating the magnitude of the 4th harmonic) are consistent across all ten fitting runs, as shown in the SI in figures S19-S23. We have not combined these results and reported the average, as is common in electrochemical literature, as the parameter variation between experiments in table 1 is larger than the calculated standard deviations (table S7). Pooling the results would obscure the experimental trends in inferred parameter values observed, which we discuss below.

For the purposes of fitting the data, normally-distributed dispersion was ultimately incorporated for two parameters, E^0 and α. E^0 was selected as the width encountered in the harmonics of the ramped data was much better described by incorporating a distribution of E^0 values. α was also selected for normally-distributed dispersion because we observed that fitting α as a point value resulted in the CMAES repeatedly returning the upper bound for α as the inferred value, regardless of where this boundary was set. This is probably attributable to the fact that k_0 values are observed to approach the reversible limit and accordingly the time series is insensitive to α. The alternative approach would be to fix the value of α at 0.5 which has the effect of in-

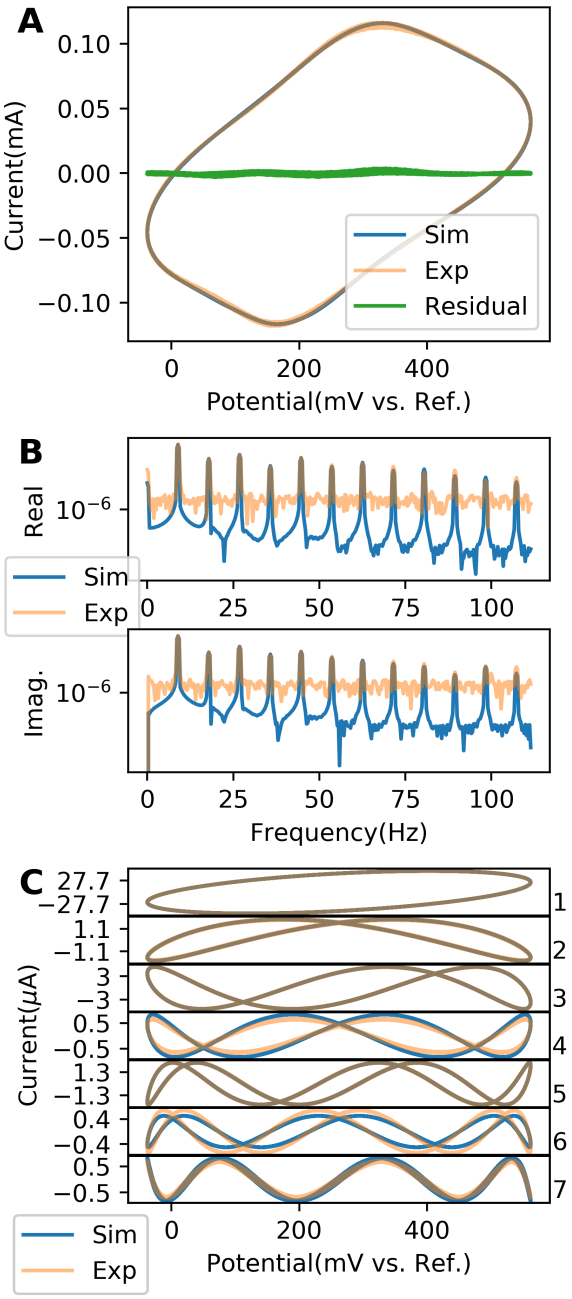

Figure 5: A comparison of simulated (blue, solid line) and experimental (orange, transparent) PSV results for oxidation of the surface confined ferrocene derivative (experiment 10) as shown in figure 1F-H., using a sinusoidal potential input with a frequency of 8.94 Hz and an amplitude of 300mV.(A) Current-voltage plot, with residual, (B) real and imaginary components of the Fourier spectrum and (C) current harmonics 1-7, plotted against voltage. The parameters used in the simulation are provided in table 1

.

creasing the inferred kinetic value to around $270s^{-1}$, but the value of the other parameters remain close to those reported in table 1. Including dispersion in the kinetic parameter did not improve the inferred fit, a phenomenon which we discuss below. In addition, the presence of dispersion in the Faradaic parameters suggests that dispersion may also be important in the simulation of current arising from double layer capacitance, although for reasons of computational feasibility, we did not explore this possibility.

Incorporation of a capacitance phase as a parameter distinct from the Faradaic phase, is essential to achieve good agreement between simulation and experiment. A difference between the two phases of about 35° is consistently observed with the ferrocene process. Modelling of the unmodified glassy carbon electrode as described in figure S25 confirms that explicitly including the capacitive phase angle is essential to obtaining a good fit. An ideal capacitor is phase-shifted by 90° relative to an input sinusoid (has the form of a cosine wave). The Faradaic current arising from a reversible process is predicted to be phase shifted by about 45° relative to the input sinusoid. However, these phase angles will be affected both by the facts that the Faradaic process is not completely reversible, and the capacitance is non-ideal. Figure S25 also demonstrates that purely capacative current has negligible current response above the second harmonic (∼3 orders of magnitude smaller than in figure 5), as with ramped FTACV.

In the parameter inference exercise based on fitting the total AC current, a global best fit of 13 parameters was achieved, as listed in table 1. On the basis of results summarised in the table 1, experiment 1 appears to be an outlier with significantly higher error estimates than found with the other experiments. This implies that a structural, orientation or other change of the surface confined ferrocene accompanies the initial oxidation-reduction cycling of potential that occurs in experiment 1. In experiments 2 to 10 the surface coverage decreases marginally

with each experiment, which can be attributed to a loss of a small quantity of ferrocene from the electrode surface accompanies each experiment. "Film-loss" is a common feature of surface confined voltammetry,[1,18] and the value of the surface concentration is corroborated to within picomolar values by analysis of DCV current in figure S27 and table S6. There are other trends in the inferred parameters, particularly in values for R_u and k_0, with results that differ by approximately $\sim$4 standard deviations (as calculated using MCMC, shown in table S7). This is presumably in response to changes in the composition of the film, and takes this form because of correlations between the parameters discussed (observed in figure 7). Significantly, the predicted distribution for α plotted in figure S24 is extremely broad, relative to the commonly "expected" α value in the range 0.4-0.6. This represents the low impact α has on PSV data when the electrode kinetics lie near to the reversible limit. The relatively wide spread of uncompensated resistance values also reflects the insensitivity of data to this parameter in highly conducting aqueous electrolyte media

It is a useful control exercise to use the parameters presented in table 1, based on analysis of PSV data, to predict the outcome of an equivalent ramped FTACV experiment. With ramped FTACV, the effect of capacitance can be minimised by focusing on the second and higher order harmonics which are essentially devoid of this non-Faradaic contribution.[15] Figure 6 provides a comparison of the second to sixth harmonics derived from a ramped FTACV experiment with the simulation output generated using the parameters inferred by the CMAES algorithm to describe PSV experiment 10. PSV Experiment 10 was chosen for this exercise as it was performed immediately prior to undertaking the ramped FTACV experiment. The fit to the ramped FTACV data is generally excellent. However, simulated and experimental plots diverge slightly. In the experimental data, the potential of the centre-point of the harmonic[6] (E_{cp}) is slightly dependent on the harmonic number, whilst the simulated E_{cp} is uniform across harmonics. It is possible to more closely replicate the experimental effect within the simulated ramped FTACV harmonics by using a distribution of E^0 values that is asymmetric, (not a "true" normal distribution). However, this experimental artefact could also be a consequence of non-idealities within the surface-confined moiety, particularly with regards to the assumption of a Langmuir isotherm, and monolayer coverage. This form of disagreement is detected in the ramped experiment because the data returned by the purely Faradaic higher order harmonics are in a format that is conducive to detecting non-ideality by visual inspection. Departures in modelling surface confined DC cyclic voltammetry based on a Butler-Volmer model of electron transfer and a Langmuir isotherm are frequently encountered in metalloenzyme studies.[36,37] However, these subtle nuances do not significantly challenge the fidelity of the parameters recovered from the PSV experiment.

The fact that the Faradaic parameters inferred by CMAES analysis of the PSV experiments also provide a good fit to ramped FTACV data indicates that the parameters inferred for the surface-linked ferrocene are all likely to be physically and chemically realistic. There is also acceptably good agreement between the parameters inferred by fitting to the 2-6th harmonics of the ramped experiment and those inferred from the PSV experiment 10 as shown in figure 6 and the fitted parameter estimates are included in table 1. The time taken to estimate parameters for ramped FTACV data was significantly greater, such that the process had to be run overnight. It should also be noted that, for the ramped FTACV fits, as the effect of capacitance is confined to the lower harmonics and these harmonics are not included in the objective function, the estimates of the parameters C_{dl} , C_{dlE1} and C_{dlE2} are unlikely to be realistic.

Analysis of the ramped FTACV data, in addition to supporting the fidelity of the computationally much less intensive PSV approach also reveal why the electrode kinetic parameter

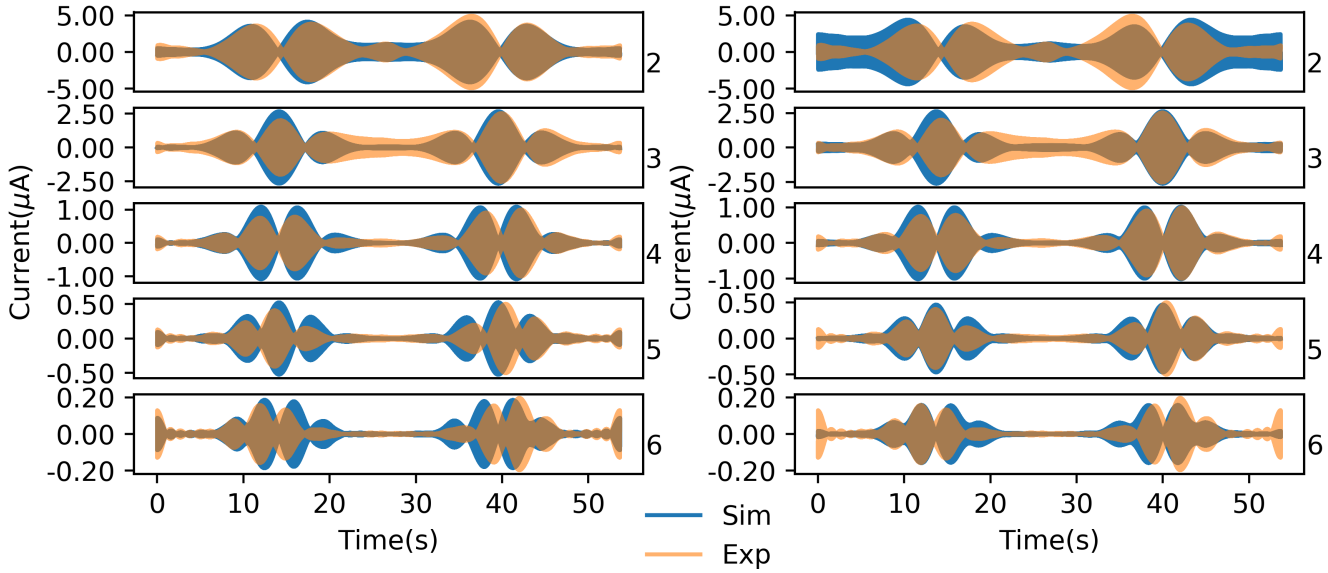

Figure 6: Comparison of harmonics 2-6 for ramped FTACV experimental data (orange) with the output of simulations (blue). The DC ramp was applied over the potential range of -0.18 to +0.62 V vs Ref at a scan rate of 29.8 mV s^{-1}. The superimposed sine wave had an amplitude of 150 mV and frequency 8.94 Hz. The simulation parameters used for the left hand side plots were as for experiment 10 in table 1, except for ω and phase, which were set at 8.885Hz and 0 rads respectively, as obtained by fitting these parameters for ramped FTACV data. The parameter values used to simulate the harmonics on the right hand side are provided in table S5. and were inferred by CMAES fitting of the 4-6th harmonics of the experimental data (shown in orange).

only needed to be fitted as a point value and not with dispersion. With the FTACV approach, the broadness of the higher order harmonics in the experimental ramped data is well-mimicked when the modelling included thermodynamic dispersion in E^0 value. An example of a fit which incorporated kinetic rather than thermodynamic dispersion is shown in figure S26 in the SI. Interestingly, metalloprotein and metalloenzyme film voltammetry analysis has often required kinetic dispersion.[36,37] The redox centres in metalloproteins and enzymes are buried within very large structure and are attached to the electrode surface with different orientations and hence distances from the electrode surface. In comparison, there should be less variation in the diazonium tethered ferrocenes with respect to orientation and distance from the GC surface. Furthermore, the k_0 value at $\sim 145s^{-1}$ approaches the reversible limit at the frequency of the experiments undertaken so that minor differences do not effect the response as significantly as would occur for slower rates of electron transfer (as shown by the difference between the $100s^{-1}$ and $150s^{-1}$ traces in the upper right of figure 3).

A measure of the confidence in parameter values for the PSV experiments, as inferred by the CMAES algorithm, was obtained by use of MCMC, the results of which are summarised in table S7. Initially an attempt was made to fit every parameter in table 1, but this did not achieve convergence of multiple MCMC chains. It was ultimately determined that the source of this problem was the $\alpha\mu$ parameter. The problem was therefore solved by removing it from the fitting process and fixing the value of $\alpha\mu$ to that inferred by the CMAES algorithm. This issue and the effect it has on the recovered distributions is discussed in the SI, with reference to figures S28 and S29. Histograms of the converged chains are provided in figure 7, where all three chains exhibit similar levels of pairwise correlation between individual parameters. Uncorrelated chains appear as a circle, and correlated chains appear as an ellipse; the narrower the ellipse the stronger the correla-

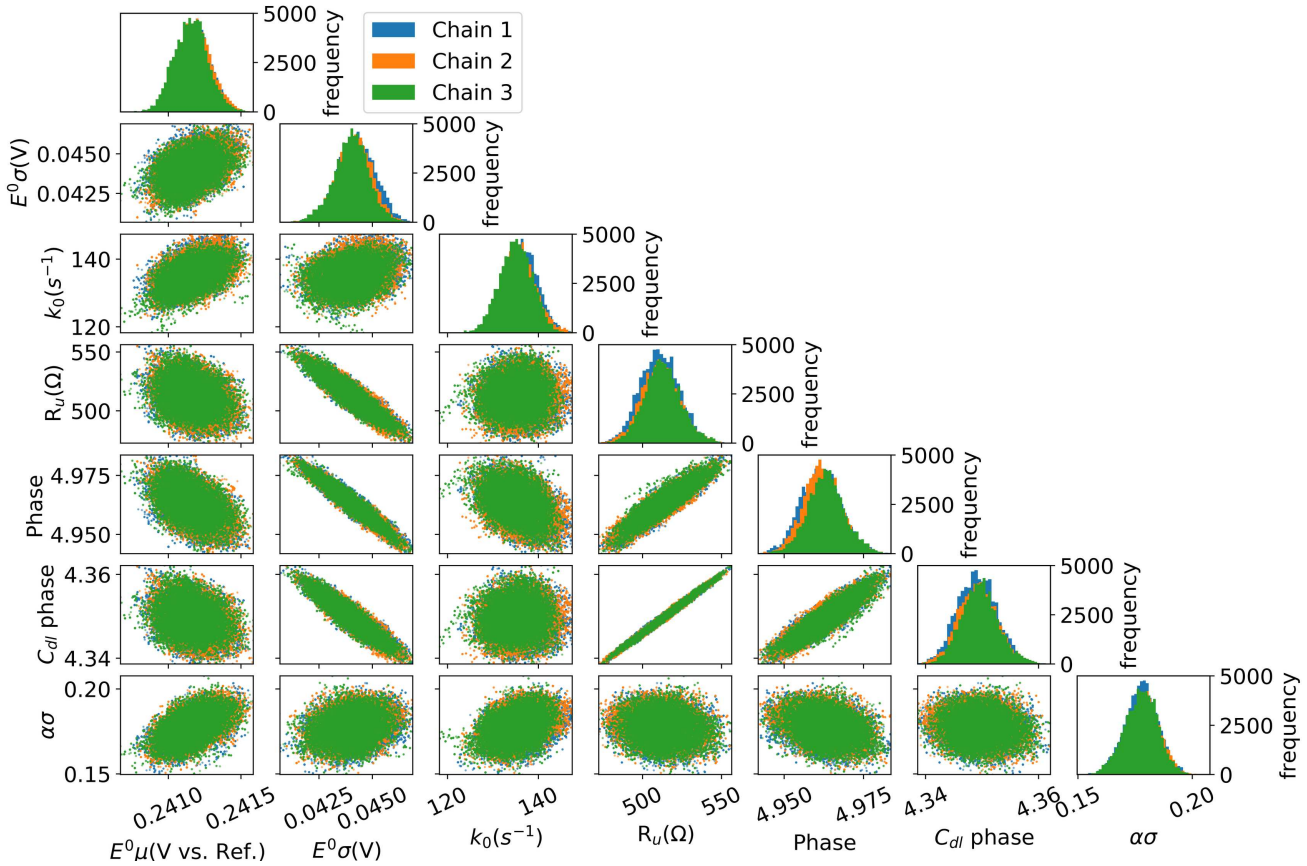

Figure 7: 2D scatter plot derived chains for individual parameters used to infer parameter distributions from PSV experiment 10, which is plotted in figure 5. The diagonal represents the 1D histograms for each parameter. The MCMC used to generated these chains was run for 100000 samples, with the first 50000 discarded as burn-in. The trace plots for each parameter are provided in figures S30-S39 in the SI.

tion. Correlation in this case indicates that the two parameters are compensating for each other. There is a strong positive correlation (i.e. an ellipse angled to the right) between the two phases, which are both also correlated with the resistance. The correlation of the phases is understood in terms of the need to maintain a constant distance between them. In addition, the fact that the uncompensated resistance is negatively correlated with the standard deviation of E^0 may be a consequence that both exert broadening effects on the PSV, as shown in figure 3. With regards to the correlation of resistance and phase, in electrical impedance, the phase of the impedance is defined by the relative sizes of the (real) resistive and (imaginary) reactance components, and so we suggest that the observed correlation is a result of this relationship. This would also explain the negative correlation between phase and thermodynamic standard deviation, through second-order correlations.

The MCMC results also establish the confidence in the mode as the inferred distributions in table S7 are almost exactly the same as those inferred by the CMAES fitting process in table 1.

Conclusions

Computational efficiency is an important issue in parameterizing surface confined voltammetry, particularly when thermodynamic and kinetic dispersion are present in the model used to simulate data. This study reveals that the computing time required to use the powerful data analysis tool of Bayesian inference to quan-

tify the 13 parameters associated with modelling of the PSV of a surface-linked diazonium based ferrocene derivative is substantially less than the more widely used, but computationally intensive dual time domain ramped form of FTACV. Although there is less confidence about the parameters inferred from PSV data (as shown by the width of the distributions inferred by MCMC in figure 4) than for ramped-FTACV data, parameters that accurately describe the PSV experiment are still accessible. The fidelity of the PSV derived parameters was demonstrated by showing that their use in simulations of the ramped FTACV experiment provided a very good fit to the data obtained with the same ferrocene modified electrode. Because of the dramatic reduction in computing time conferred by using PSV, the impact of multiple dispersed parameters can be addressed with a desktop computer on the timescale of an hour. In contrast it is computationally impractical to solve the inverse problem with the ramped experiment without access to a machine with considerably more computing power. The dramatic reduction in computing time, which is the core benefit of PSV relative to ramped FTACV, has therefore facilitated a much more thorough exploration of the surface confined process. However, inspection of the ramped data set, which is visually much more informative can still be used to assist in decisions about which form of parameter dispersion to include in the model.

Supporting Information Available

Additional mathematical details, including a description of the model and further details about dispersion. Detailed experimental methods including synthetic methods. Supplemental simulation details and results from the fitting process.

Acknowledgement HOLL gratefully acknowledges funding from the EPSRC and BBSRC Centre for Doctoral Training in Synthetic Biology (grant EP/L016494/1). NDJY was supported by the BBSRC (studentship BB/M011151/1), ARH by the York Department of Chemsistry summer research bursary and MJR and DJG gratefully acknowledge support from the EPSRC Centres for Doctoral Training Programme (EP/S024093/1). JZ, AMB, DJG and AP thank the Australian Research Council for financial support that facilitated this international collaboration under the auspices of Discovery Program Grant DP170101535.

References

(1) Adamson, H.; Bond, A. M.; Parkin, A. *Chem. Commun.* **2017**, *53*, 9519–9533.

(2) Yan, M.; Kawamata, Y.; Baran, P. S. *Chem. rev.* **2017**, *117*, 13230–13319.

(3) Braun, A. *Electrochemical energy systems: foundations, energy storage and conversion*; Walter de Gruyter GmbH & Co KG, 2018.

(4) Britz, D.; Strutwolf, J. *Digital simulation in electrochemistry*; Springer, 2005; Vol. 666.

(5) Gavaghan, D. J.; Bond, A. M. *J. Electroanal. Chem.* **2000**, *480*, 133–149.

(6) Adamson, H.; Robinson, M.; Wright, J. J.; Flanagan, L. A.; Walton, J.; Elton, D.; Gavaghan, D. J.; Bond, A. M.; Roessler, M. M.; Parkin, A. *J. Am. Chem. Soc.* **2017**, *139*, 10677–10686.

(7) Allen, J. B.; Larry, R. F. *Electrochemical methods fundamentals and applications*; John Wiley & Sons, 2001.

(8) Robinson, M.; Ounnunkad, K.; Zhang, J.; Gavaghan, D.; Bond, A. M. *ChemElectroChem* **2019**, *6*, 5499–5510.

(9) Bieniasz, L.; Speiser, B. *J. Electroanal. Chem.* **1998**, *458*, 209–229.

(10) Gavaghan, D. J.; Cooper, J.; Daly, A. C.; Gill, C.; Gillow, K.; Robinson, M.; Simonov, A. N.; Zhang, J.; Bond, A. M. *ChemElectroChem* **2018**, *5*, 917–935.

(11) Bond, A. M. *Journal of Solid State Electrochemistry* **2017**, 2041–2050.

(12) Robinson, M.; Simonov, A. N.; Zhang, J.; Bond, A. M.; Gavaghan, D. *Anal. Chem.* **2018**, *91*, 1944–1953.

(13) Armstrong, N.; Hibbert, D. *Chemom. Intell. Lab. Syst.* **2009**, *97*, 194–210.

(14) Samuel, J. J. Empirical models for cyclic voltammograms. Ph.D. thesis, University of Southhampton, 2010.

(15) Bond, A. M.; Duffy, N. W.; Guo, S.-X.; Zhang, J.; Elton, D. *Anal. Chem.* **2005**,

(16) Bond, A. M.; Elton, D.; Guo, S.-X.; Kennedy, G. F.; Mashkina, E.; Simonov, A. N.; Zhang, J. *Electrochem. Commun.* **2015**, *57*, 78–83.

(17) Guo, S.-X.; Bond, A. M.; Zhang, J. *Review of Polarography* **2015**, *61*, 21–32.

(18) Léger, C.; Bertrand, P. *Chem. Rev.* **2008**, *108*, 2379–2438.

(19) Evans, R. M.; Siritanaratkul, B.; Megarity, C. F.; Pandey, K.; Esterle, T. F.; Badiani, S.; Armstrong, F. A. *Chem. Soc. Rev.* **2019**, *48*, 2039–2052.

(20) Morris, G. P.; Baker, R. E.; Gillow, K.; Davis, J. J.; Gavaghan, D. J.; Bond, A. M. *Langmuir* **2015**, *31*, 4996–5004.

(21) Adamson, H.; Robinson, M.; Bond, P. S.; Soboh, B.; Gillow, K.; Simonov, A. N.; Elton, D. M.; Bond, A. M.; Sawers, R. G.; Gavaghan, D. J. *Anal. Chem.* **2017**, *89*, 1565–1573.

(22) Heyrovskỳ, J.; Kůta, J. *Principles of Polarography*; Academic Press: New York, 1966.

(23) Bell, C. G.; Anastassiou, C. A.; O'Hare, D.; Parker, K. H.; Siggers, J. H. *Electrochim. Acta* **2011**, *56*, 7569–7579.

(24) Long, J. T.; Weber, S. G. *Electroanalysis* **1992**, *4*, 429–437.

(25) Hebert, N. E.; Kuhr, W. G.; Brazill, S. A. *Anal. Chem.* **2003**, *75*, 3301–3307.

(26) Brazill, S. A.; Bender, S. E.; Hebert, N. E.; Cullison, J. K.; Kristensen, E. W.; Kuhr, W. G. *J. Electroanal. Chem.* **2002**, *531*, 119–132.

(27) Guo, Z.; Lin, X. *Anal. Lett.* **2005**, *38*, 1007–1017.

(28) Sher, A. A.; Bond, A. M.; Gavaghan, D. J.; Gillow, K.; Duffy, N. W.; Guo, S.-X.; Zhang, J. *Electroanalysis* **2005**, *17*, 1450–1462.

(29) Salverda, J. M.; Patil, A. V.; Mizzon, G.; Kuznetsova, S.; Zauner, G.; Akkilic, N.; Canters, G. W.; Davis, J. J.; Heering, H. A.; Aartsma, T. J. *Angew. Chem. Int. Ed.* **2010**, *49*, 5776–5779.

(30) Hauquier, F.; Debou, N.; Palacin, S.; Jousselme, B. *J. Electroanal. Chem.* **2012**, *677*, 127–132.

(31) Yates, N. D. J.; Dowsett, M. R.; Bentley, P.; Dickenson-Fogg, J. A.; Pratt, A.; Blanford, C. F.; Fascione, M.; Parkin, A. *Langmuir* **2019**, 5654–5664.

(32) Robinson, M.; Ounnunkad, K.; Zhang, J.; Gavaghan, D.; Bond, A. *ChemElectroChem* **2018**, *5*, 3771–3785.

(33) Hansen, N. *arXiv preprint arXiv:1604.00772* **2016**,

(34) Clerx, M.; Robinson, M.; Lambert, B.; Lei, C. L.; Ghosh, S.; Mirams, G. R.; Gavaghan, D. J. *J. Open Res. Softw.* **2019**,

(35) Gelman, A.; Stern, H. S.; Carlin, J. B.; Dunson, D. B.; Vehtari, A.; Rubin, D. B. *Bayesian data analysis*; Chapman and Hall/CRC, 2013.

(36) Fourmond, V.; Léger, C. *Curr. Opin. Electrochem.* **2017**, *1*, 110–120.

(37) Hexter, S. V.; Esterle, T. F.; Armstrong, F. A. *Phys. Chem. Chem. Phys* **2014**, *16*, 11822–11833.

Graphical TOC Entry