

Leveraging e-Science infrastructure for electrochemical research

BY TOM PEACHEY¹, ELENA MASHKINA², CHONG-YONG LEE²,
COLIN ENTICOTT¹, DAVID ABRAMSON^{1,*}, ALAN M. BOND², DARRELL ELTON³,
DAVID J. GAVAGHAN⁴, GARETH P. STEVENSON⁴ AND GARETH F. KENNEDY²

¹*Faculty of Information Technology, and* ²*School of Chemistry,
Monash University, Clayton, VIC 3800, Australia*

³*Department of Electronic Engineering, Latrobe University,
Bundoora, VIC 3083, Australia*

⁴*Computing Laboratory, University of Oxford, Wolfson Building,
Park Road, Oxford OX1 3QD, UK*

As in many scientific disciplines, modern chemistry involves a mix of experimentation and computer-supported theory. Historically, these skills have been provided by different groups, and range from traditional ‘wet’ laboratory science to advanced numerical simulation. Increasingly, progress is made by global collaborations, in which new theory may be developed in one part of the world and applied and tested in the laboratory elsewhere. e-Science, or cyber-infrastructure, underpins such collaborations by providing a unified platform for accessing scientific instruments, computers and data archives, and collaboration tools. In this paper we discuss the application of advanced e-Science software tools to electrochemistry research performed in three different laboratories – two at Monash University in Australia and one at the University of Oxford in the UK. We show that software tools that were originally developed for a range of application domains can be applied to electrochemical problems, in particular Fourier voltammetry. Moreover, we show that, by replacing ad-hoc manual processes with e-Science tools, we obtain more accurate solutions automatically.

Keywords: scientific workflows; Fourier voltammetry; parameter sweeps electrochemistry

1. Introduction

The interplay between observation and computation has long been a central part of scientific discovery and science-based design. With the advent of the digital computer, these computations have become more complex, giving detailed predictions that can not only model existing data, but also be used prognostically to explore new scenarios. As a result, computational models have spread to most areas of science and engineering and even to social sciences [1]. In practice, multiple runs are often required to validate that the outcomes agree with those of the real system. In particular, many models involve parameters for which

*Author for correspondence (david.abramson@infotech.monash.edu.au).

One contribution of 12 to a Theme Issue ‘e-Science: novel research, new science and enduring impact’.

the values are unknown, so a calibration phase is required to select these. In certain applications, the determination of these parameters is the main aim of the work; values are chosen to give the best fit between experimental data and model outputs. This paper presents just such an application in the field of electrochemistry.

Dynamic electrochemical techniques can measure the response of a chemical system when an electrical input signal is applied. In the case of voltammetry, the perturbation is a potential V , and the response is a current as a function of time [2,3]. To date, voltammetry has been widely applied to quantify the kinetics and thermodynamics associated with electron transfer processes. These are of interest because of their ubiquity and essential role in many physical, chemical and biological systems.

In this study, two collaborating groups, one in the School of Chemistry at Monash University (Australia) and one at the University of Oxford (UK), have been investigating a computational model of the one-electron transfer process in voltammetry. Specifically, the work models the Butler–Volmer equation [2] for the chemical reaction: $\text{Red} \rightleftharpoons \text{Ox} + \text{e}^-$. This model predicts the net current flowing at the working electrode, as a function of the electron transfer rate constant, traditionally denoted by k_0 ; determination of that constant is the prime purpose of this work. Other terms in the model are either known constants or can be directly measured during an experiment, with the exception of the uncompensated resistance R_u and the electrochemical double-layer capacitance function $C_{dl}(V)$.

Experimentation on real-world systems is known to be typically iterative. Unforeseen problems arise, instruments may not act as expected, results seem inconsistent and even the research questions being addressed may need modification. A leading authority on the design of experiments, G. E. Box, recommends an incremental approach, starting with small preliminary trials [4] rather than a large all-inclusive experiment. In the same way, our experience is that computational experiments also require this incremental methodology. The work described below shows such early confusion and dead ends. Consequently, tools that are used to assist computer experiments should be generic, able to tackle a variety of situations and flexible, easily configured to a new experiment.

Over a number of years, we have built such a family of generic e-Science tools, called Nimrod, that support the experimental approach discussed above. Nimrod allows scientists to integrate physical experiments with computational ones and supports the exploration of physical parameters that cannot be measured in the laboratory. It leverages the substantial investment in e-Science infrastructure that has occurred in many countries, allowing users to run computations on a range of systems. Importantly, though, it shields users from much of the complexity in the infrastructure, allowing them to focus on their science without concern for the many complexities in running real infrastructure.

This paper highlights our success in applying some generic e-Science infrastructure to a real-world problem in electrochemistry and also highlights the value of such multi-disciplinary, multi-site, collaborations. In §2, we discuss the background to the voltammetry techniques and in §3 we discuss the Nimrod tool family. Section 4 then shows how we have applied the Nimrod tools to voltammetry and, in particular, we highlight the iterative nature of the experimental and computational work. We show that the tools are powerful

enough to allow us to make progress on the science without being distracted by the infrastructure. The outcome of the work is that we have achieved new scientific insights that would have been difficult without the e-Science infrastructure.

2. Voltammetry

Voltammetry is a chemical technique for determining properties of an analyte using electrochemical cells. The analyte under investigation is either in solution in the cell or is confined to the working electrode. The cell potential is varied in some way and the resulting current is measured. The many variants of the technique differ mainly in the waveform of the applied potential. Information from voltammetry experiments is obtained by comparing the current measured with theoretical models. In particular, the Butler–Volmer equation relates applied potential to the electron transfer at the working electrode, and the diffusion equation combines with the potential and cell geometry to predict the Faradaic current. Our work uses computational models based on these theories.

Various codes implementing these standard voltammetry models are in current use. We employed two of these: MECSIM (Monash Electrochemistry Simulator) is a Fortran implementation developed by Kennedy, and the other implementation used is a Matlab program, built by Stevenson and Gavaghan, all co-authors of this paper.

The work described here employs large-amplitude Fourier-transform (FT) alternating current (AC) voltammetry. Here the potential that is applied consists of two components. One is a linearly increasing, followed by a linearly decreasing, ramp function as used in conventional cyclic voltammetry. The other component is a sinusoidal function with frequency ω . Since the system is nonlinear, the response current shows a power spectrum, with peaks at integral multiples of ω . The current component that corresponds to the applied frequency band (ω) is designated the fundamental harmonic, while the second, third, fourth, fifth, etc. harmonics are associated with the response at frequencies of 2ω , 3ω , 4ω , 5ω , respectively. The component associated with frequencies around zero is termed the direct current (DC) component.

Figure 1 represents the data-processing strategy used for large-amplitude FT AC voltammetry for a single sinusoidal frequency signal. The total current, as a function of time, is subjected to a Fourier transform to yield the power spectrum. Specifically, a discrete FT algorithm converts the time-domain data to the frequency domain and is presented in the form of a power spectrum. The relevant region (DC or harmonic of interest) in the power spectrum is then selected, and the power of all other frequencies is set to zero, thereby removing the higher-order frequency components. This is followed by an inverse discrete FT to give the required component. The DC component resembles a conventional cyclic voltammogram, but differs because of the presence of the sinusoidal addition.

The traditional procedure for the determination of model parameters in large-amplitude FT AC voltammetry uses a computer-controlled physical experiment to obtain voltammetry data for a given analyte. Then, a computational model is executed repeatedly (manually) with various parameters, and results are evaluated by observing the differences between waveform envelopes for the physical and simulated systems in the so-called ‘heuristic approach’ [5,6].

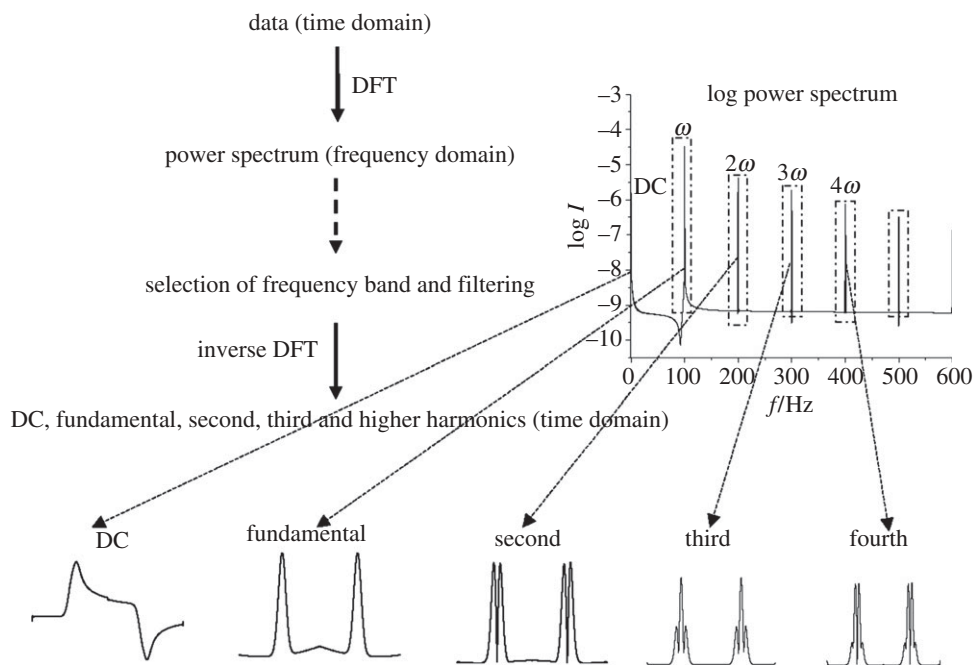


Figure 1. Data processing strategy in FT voltammetry.

Currently models are executed one at a time, and informal judgement is used to choose the next parameter set. This manual procedure has shown promise in predicting electrochemical parameters: the electron transfer rate constant k_0 , charge transfer coefficient α , uncompensated resistance R_u and the electrochemical double-layer capacitance function C_{dl} . Other parameters, such as temperature, concentration, diffusion coefficient, mid-point potential and effective electrode area, are known values from the literature, or recorded with the physical experiment, or may be determined from independent experiments. In particular, the electrode area A may be found from the Randles–Sevcík equation and the formal electrode potential E° recovered from the DC component of the measured current. However, the manual procedure involves significant human effort. It is also hard to repeat because the operator needs to manually record the state of the experiment and decide what values to choose in an ad-hoc way. In this paper, we will show that it is possible to automate this process using advanced generic e-Science tools, obtaining better and more robust scientific outcomes, with reduced manual effort. In addition, we show that it is possible to leverage generic e-Science infrastructure such as computers and data storage systems.

3. The Nimrod toolkit

Computational models often involve a ‘workflow’, a sequence of discrete tasks with information flowing between them. However, a single run of a model is rarely sufficient; the experimenter will want to explore the parameter space

to determine how variation of the input parameters affects the outputs. Then there will be a higher-level workflow made up of the individual runs. In recent years, the MeSSAGE Lab team at Monash University have identified common patterns in computational workflows and have built a set of tools to assist in such work, thereby producing the Nimrod toolset. This section will describe some of these tools.

The original Nimrod tools are UNIX applications that may be invoked at the command line or run via the Nimrod portal. These require a workflow description written in a simple declarative language (for Nimrod/G and Nimrod/O), stored in a file called a plan file; such files may be produced manually or automatically generated by the Nimrod portal. The extensive commonality between plan files for different tools enables the user to quickly change between them. A recent addition to the toolkit, Nimrod/K, provides similar functionality via the KEPLER workflow engine. Nimrod/K workflows are stored as KEPLER XML (eXtensible Markup Language) files and are typically manipulated using a graphical user interface.

(a) *Nimrod/G*

Nimrod/G is a parameter sweep tool. Each input parameter is assigned a set of values. The tool generates all possible combinations of these and executes the jobs corresponding to each combination. For most experiments, these computations are independent, so Nimrod/G will run them concurrently insofar as computational resources allow. Early experiments with Nimrod/G were run on a cluster [7]. Later, the use of computational grids [8] and clouds [9] allowed very large experiments. What is important is that the user is largely shielded from the differences in computational infrastructure.

Nimrod/G uses whatever resources are currently available, and jobs are automatically transferred from unresponsive resources. The user may specify resources via the Nimrod portal. These may be added or deleted while an experiment is running.

(b) *Nimrod/O*

Many users of computational models seek to optimize some aspect of the model output. Engineering designs are an obvious area for such studies [10]. Also common are the inverse problems, finding input parameters that yield model outputs that agree closely with those of real-world systems. Such problems may be treated as an optimization, minimizing some measure of the difference between the model outputs and those of the real system. The parameter estimation in large-amplitude FT AC voltammetry is an example of this work.

Nimrod/O [11] performs such experiments. It supplies a rich collection of optimization algorithms, interfacing with Nimrod/G for job execution. Most algorithms are iterative, generating batches of jobs for evaluation (in parallel if multiple processors are available) and using the results to decide the next batch. So the workflow is cyclic in a Nimrod/O experiment. Design of a Nimrod/O experiment requires specification of which algorithm(s) to use, their settings and criteria for termination of the search.

(c) Nimrod/K

Many tools have been developed to facilitate scientific workflows [12]. These encapsulate the tasks and handle the flow of information between them. In particular, graphical workflow engines, such as KEPLER [13], represent the tasks as screen icons and the information flow by connecting arrows. The user may thus create novel workflows just by selecting the tasks (called *actors* in KEPLER) from a palette and connecting them. Nimrod/K [14] is a version of KEPLER that incorporates Nimrod functionality such as parameter sweep generation and distributed execution. A version of Nimrod/O is also included under the name Nimrod/OK [15].

4. Experimental results and discussion

The aim of these investigations was to determine whether the process of parameter fitting for FT voltammetry could be automated using the Nimrod tools and, more ambitiously, to see whether Nimrod could expedite the process, yield better solutions and, in general, provide deeper insights into the chemistry. Here we demonstrate the iterative and incremental nature of computer experiments and show how the flexibility of the tools facilitates such work. Notably, we started out using only one particular tool (Nimrod/O) but moved on to other members of the family as the work developed (Nimrod/G and Nimrod/K).

Initially, we tested the solution on artificial data to determine whether our overall approach was correct. We then moved to real data gathered from experiments; the details of these experiments and associated chemistry are discussed in appendix A.

(a) Tests on artificial data

The initial computational experiments were conducted using the Matlab voltammetry code of Stevenson and Gavaghan. For this, parameters are read from a Matlab M file, and these are automatically placed in that file under Nimrod. A model of a large-amplitude FT AC voltammetry experiment, based on parameters relevant to the one-electron reduction ($\text{Cu(II)} + \text{e}^- \rightleftharpoons \text{Cu(I)}$) confined azurin process, was prepared using an estimate of reasonable values for the unknown parameters, namely $k_0 = 1000 \text{ (s}^{-1}\text{)}$, $R_u = 500 \text{ (}\Omega\text{)}$, $C_{dl} = 2.0 \text{ (}\mu\text{F cm}^{-2}\text{)}$ and $\alpha = 0.50$ [5]. This model was used to generate 2^{12} data points for the output current, and these were stored in a file for later processing.

A Nimrod/O experiment was performed to determine whether the parameters k_0 , R_u and C_{dl} could be recovered from these surrogate data, treating the system as a classic inverse problem. Extra code was added to the Matlab model to read the surrogate data and compute the sum of squares of the differences between model and experimental data for the full output signals; this will be called metric M1. (The sum of squares is a fairly conventional way of measuring the difference between a signal and some computational output.)

Figure 2 shows the plan file used by the Nimrod/O tool. The three parameters varied and their ranges are defined by the parameter statements. The task section defines the operations required for the execution of a single job. In this case, requisite files are copied to the location where execution will be performed. Then

```
parameter rate float range from 100 to 2000
parameter res float range from 0 to 1000
parameter cap float range from 1e-6 to 10e-6

task main
  copy root:*.m node:.
  copy root:dim_params.sk node:.
  copy root:exptal_current.txt node:.
  node:substitute dim_params.sk dim_params.m
  node:execute matlab_batchjob main
  copy node:results.dat root:output.$jobname
endtask

method simplex
  starts 5
    starting points random
    tolerance 5e-12
  endstarts
endmethod
```

Figure 2. Nimrod/O plan file for parameter fitting.

Table 1. Results of five optimizations fitting surrogate data.

starting parameters			optimal parameters			metric 1 (10^{-18} A^2)
$k_0 \text{ (s}^{-1}\text{)}$	$R_u \text{ (}\Omega\text{)}$	$C_{dl} \text{ (F cm}^{-2}\text{)}$	$k_0 \text{ (s}^{-1}\text{)}$	$R_u \text{ (}\Omega\text{)}$	$C_{dl} \text{ (}\mu\text{F cm}^{-2}\text{)}$	
1836	704.7	9.048	2000	1000	1.0	86.301
1667	90.4	3.859	1000	500	2.0	20.136
1211	816.0	5.084	1000	500	2.0	20.112
1586	222.9	2.723	1000	500	2.0	20.112
804	541.4	1.401	1000	500	2.0	20.112

the *substitute* command instantiates the parameters in the Matlab file. After execution, the results file is copied back and given a unique name. This file must contain the objective function for the optimization; in this case, the M1 metric. The final section specifies which optimization algorithm to use, how many independent searches and the criteria for terminating a search.

This experiment used five separate optimizations, using the simplex algorithm of Nelder & Mead [16], and was run on an Intel Xeon E5310, 1.6 GHz, providing a total of eight processor cores. The starting points for the searches were randomly selected in the domain $100 \leq k_0 \leq 2000$, $0 \leq R_u \leq 1000$ and $1.0 \leq C_{dl} \leq 10.0$. Results, shown in table 1, illustrate the importance of multiple searches. Four searches recovered the known values of the parameters, while the other search terminated at a local minimum on the boundary of the search space.

This experiment confirmed that the basic approach worked – we had replaced the ad-hoc manual process for finding parameters with a fully automated one using one of the Nimrod tools family. The nonlinear behaviour of the system

meant that we needed to do multiple runs. Nimrod allows us to exploit eight processor cores to do the searches in parallel, without any changes to the Matlab-based computational model, a significant advantage. (We have not reported speedups here since that is not the focus of this work and has been discussed in many other publications about Nimrod [8,9,14].)

(b) *Fitting real data*

Encouraged by these results, we attempted more realistic experiments. A large-amplitude FT AC voltammetry experiment was performed in aqueous 0.5 M KCl electrolyte medium using potassium ferricyanide ($\text{K}_3[\text{Fe}(\text{CN})_6]$), generating 2^{17} data points. In this case, the $[\text{Fe}(\text{CN})_6]^{3-} + \text{e}^- \rightleftharpoons [\text{Fe}(\text{CN})_6]^{4-}$ reduction process at a glassy carbon electrode was of interest. Again Nimrod/O was used to estimate the parameters and again the parameters varied were k_0 , R_u and C_{dl} . The value 0.5 was assumed for α because this is the most likely value for simple electron transfer. Various metrics were used: M1, as described above, and a number of other metrics based on the differences for single harmonic components.

However, unlike the earlier experiment, none of these gave satisfactory results. Typically the optimal parameters found were the limiting values of the search space itself, indicating that the search terminated when it hit the boundary. In particular, k_0 was zero, implying that the optimal output current was identically zero, clearly an incorrect result. Figure 3 shows the third harmonics for experimental (red) and model (blue) current. This graph shows a good fit between the experiment and the model, but the metrics we had computed actually reported a bad value. In order to try and understand the discrepancy, we produced a higher-resolution graph in figure 4, which reveals the cause of the problem. The experimental and model sinusoids are out of phase by 90° . The phase applied in the physical experiment is not known when building the computational model, so the M1 metric gives a spurious measure of parameter fit.

To remedy this problem, we developed a new metric. This time, the experimental and model outputs were filtered to produce the various harmonics. Then a spline envelope was computed for these waveforms and a metric computed using the sums of squares of the deviations between the two envelopes. The value thus obtained from the n th harmonic is denoted by En . (The phase information could also have been removed in the Fourier domain, but this was not attempted.)

Using the new metric En , Nimrod/O searches converged to interior points in the search space, suggesting that we had in fact fixed the cause of the error. However, multiple optimizations showed great variation in the results obtained, suggesting that there were actually a large number of suitable solutions. The changes required here involved writing new metrics, and this could be done with some changes to the Matlab code used to implement M1.

(c) *Changing the computational modelling software*

At this time, we switched to a Fortran version of the MECSIM code, mostly because it was significantly faster than the Matlab version. To compute a metric comparing model results with experimental ones, a Matlab code was built, so in this case the workflow for a single evaluation consisted of the MECSIM execution followed by a Matlab one. For these experiments, a more complex capacitance function was used, defined by a second-order polynomial with three parameters,

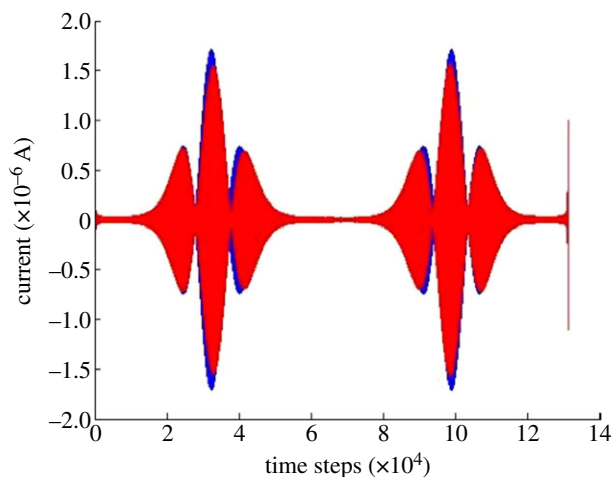


Figure 3. Comparison of third harmonics of experimental (red) and model led (blue) current for the $[\text{Fe}(\text{CN})_6]^{3-} + \text{e}^- \rightleftharpoons [\text{Fe}(\text{CN})_6]^{4-}$ process.

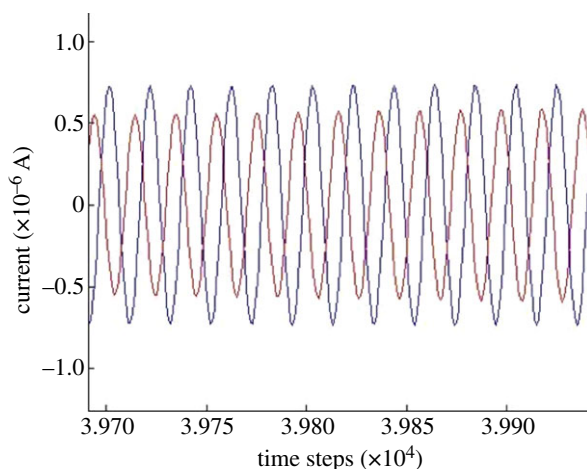


Figure 4. Enlargement of a portion of the data from figure 3.

denoted by C_0 , C_1 and C_2 [17]. These, together with the electron transfer rate k_0 and the uncompensated resistance R_u , constitute a six-parameter space to be searched. The charge transfer coefficient α was still assumed to be 0.5.

Even though we could have continued to use the original Nimrod/O plan file, the added complication of running a Fortran code for the model, followed by a Matlab-based objective function, caused us to shift to the Nimrod/K tool. This allowed us to build a more complex pipeline of computations and, moreover, we could represent the flow using a graphical programming environment. Thus, while the previous implementation of the workflow used a simple text-based declarative programming language, we now used a graphical display of the components, supported by the KEPLER workflow engine [13].

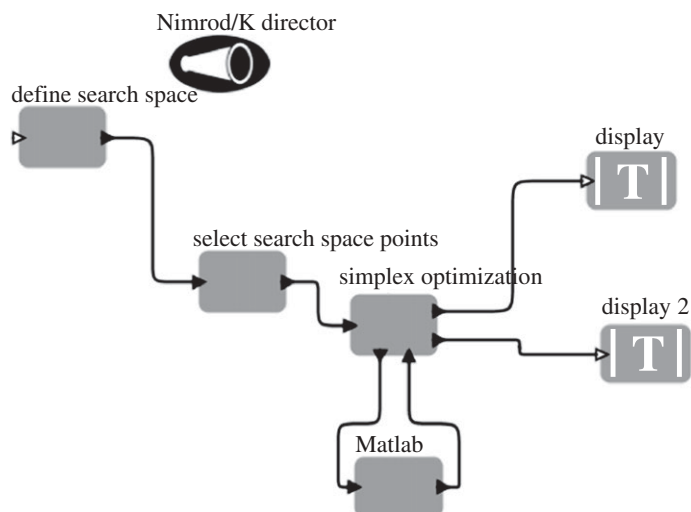


Figure 5. Nimrod/K-based workflow for simplex optimization.

Figure 5 shows such a workflow. The parameters and their domains are entered in the actor ‘define search space’. Optimization searches require starting points; the strategy for selecting these is supplied in ‘select search space points’. The number of starting points determines the number of searches. ‘Simplex optimization’ implements the simplex algorithm with settings and convergence criteria as entered. When the workflow executes, this actor generates batches of jobs for execution by the Matlab code interfaced by the ‘Matlab’ actor. On convergence, each optimization result is forwarded to the ‘display’ actor, and statistical summaries to ‘display 2’.

In spite of a new computation model, experiments still gave inconsistent results.

(d) A two-stage optimization

At this stage, it was decided to adopt a methodology that would mimic that used in manual parameter fitting. This was a two-stage procedure. Optimizations over all six parameters were performed using the fundamental harmonic metric E1. That harmonic is sensitive to the capacitance function [5,17], so the capacitance parameters were selected to be those found in the best of these optimizations (the one with the minimum E1 value). With these values fixed, multiple optimizations were performed to determine the remaining parameters k_0 and R_u using the envelope metric for a higher harmonic, known to be more sensitive to those parameters. The first experiments with this new approach were conducted with ferrocene (Fc), this time not in aqueous medium but in the much higher-resistance organic solvent acetonitrile containing 0.1 M Bu_4NPF_6 as the electrolyte. The electron transfer process of interest was the one-electron oxidation $\text{Fc}^0 \rightleftharpoons \text{Fc}^+ + \text{e}^-$, again at a glassy carbon electrode.

Figure 6a shows the Nimrod/K workflow that will perform this two-stage process. The ‘voltammetry’ actors are composite, made up of the MECSim simulation followed by Matlab code that computes the envelope metric; in the case

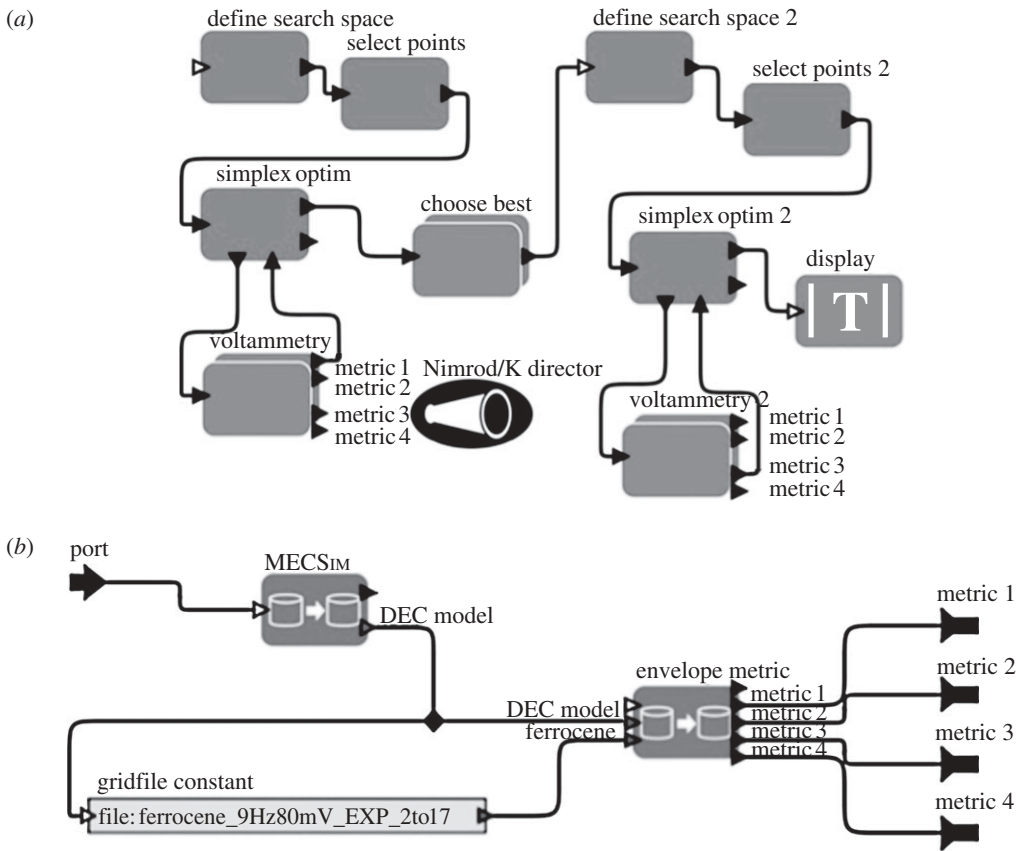


Figure 6. (a) A two-stage workflow for voltammetry parameter fitting. (b) Expansion of the ‘voltammetry’ composite actor.

Table 2. Parameter values for oxidation of ferrocene obtained using the experiment depicted in figure 6.

starting parameters		optimal parameters		metric E3 (10^{-9} A ²)
k_0 (cm s ⁻¹)	R_u (Ω)	k_0 (cm s ⁻¹)	R_u (Ω)	
0.9135	155.7	0.8558	149.5	1.27068
0.8942	173.7	0.9768	150.3	1.27417
0.0904	97.65	0.4453	143.2	1.26573
0.5850	172.4	0.7467	148.4	1.26815
0.4538	167.3	0.5362	144.9	1.26393

shown, comparison is made with physical data for ferrocene. Figure 6b displays the constituent components of these actors, showing the execution of the MECSIM code followed by the ‘envelope metric’.

Table 2 gives the results for five optimizations using the two-stage optimization. The results show that, although the resistance values are somewhat consistent, the rate constant (k_0) varies dramatically.

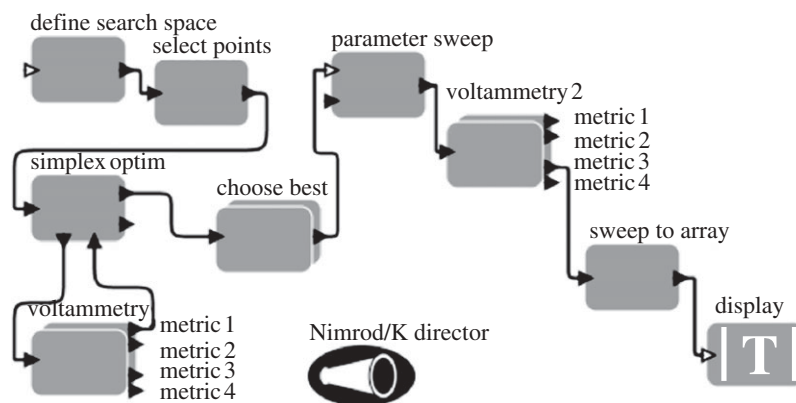
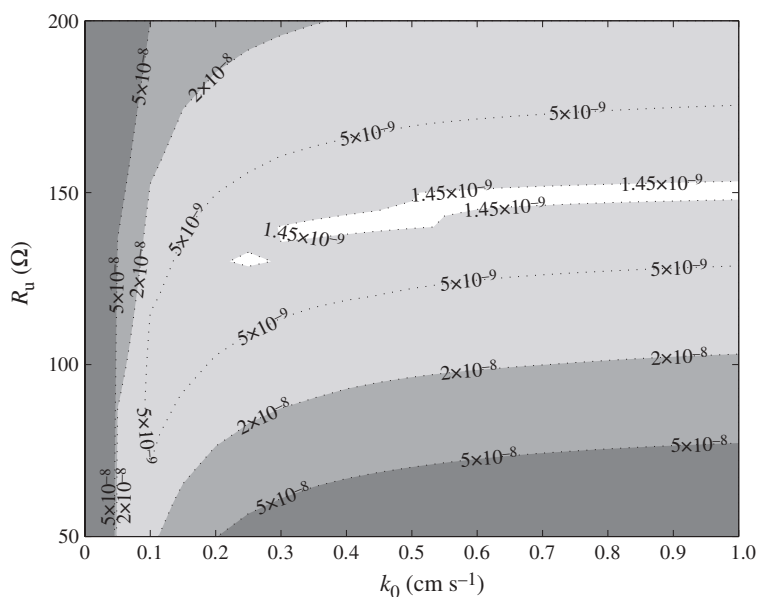


Figure 7. An optimization followed by a sweep.

Figure 8. Contour map for the metric E3 against k_0 and R_u for the oxidation of ferrocene.

(e) Optimization followed by a sweep

To further illuminate these results, a new experiment was conducted. The first stage remained the same, but the second performed a sweep over k_0 and R_u , rather than attempting to find optima. Modification of the experiment in this way is a simple task in Nimrod/K: a matter of replacing one part of the workflow with another. Figure 7 shows the new experiment; the second stage of the workflow is now simpler, as a sweep consists of a single batch of jobs.

Figure 8 is a contour map produced from these results, in this case that generated using the third harmonic metric E3. Maps generated using E2, E3 and E4 are all quite similar. They explain the variation of results from the optimization; the local minima occur in various bands (white in the diagram), all with similar values for resistance but with great variation in the rate.

Table 3. Parameter values for reduction of ferricyanide using a two-stage optimization.

starting parameters		optimal parameters		metric E3 (10^{-10} A ²)
k_0 (cm s ⁻¹)	R_u (Ω)	k_0 (cm s ⁻¹)	R_u (Ω)	
0.01827	70.47	0.01103	97.5	5.69627
0.01789	82.48	0.01114	100.0	5.71050
0.00181	31.77	0.00947	0.986	5.59123
0.01170	81.6	0.01010	40.04	5.44034
0.00908	78.19	0.01019	46.65	5.44054

This experiment yields a wide range of values for k_0 for ferrocene, with the lower limit of ca. 0.25 cm s⁻¹. The bound $k_0 > 0.25$ cm s⁻¹ is much larger than the ones proposed in the literature by Tsierkezos & Ritter [18] and Xiao *et al.* [19] and were measured by the DC cyclic voltammetry method. This indicates that a higher detection limit is accessible in the large-amplitude FT AC voltammetric technique.

(f) Results for potassium ferricyanide

We now returned to experiments with the potassium ferricyanide data using the two-stage optimization procedure. This enabled us to use the same computer experiments for parameter fitting, with just the experimental results file changed. Table 3 shows the results of the parameter optimizations. In this case, the estimates for the electron transfer rate are consistent, but the uncompensated resistance shows no pattern at all. Again a sweep was used to prepare contour maps, which explained these results. Figure 9 is the map for the third harmonic, again typical of all the maps. It reveals that the metric used is insensitive to the resistance and confirms the sharp estimate for k_0 .

This experiment consistently yields a k_0 value around 0.010 cm s⁻¹, with R_u any value below 100 Ω . This quasi-reversibility rate constant value is expected for an inner-sphere process [2]. The k_0 values at glassy carbon electrode surfaces are very dependent on the treatment history and the method of manufacture [20]. Moreover, the presence of 0.5 M KCl as supporting electrolyte significantly minimizes the solution resistance, resulting in an insignificant R_u value. The employed harmonics are not sensitive to this level of uncompensated resistance and, therefore, the optimization procedure is not able to yield a localized estimate of R_u .

Until now, the experiments had assumed that charge transfer coefficient α was equal to 0.50. In order to test this, a final experiment on the ferricyanide data was conducted with R_u fixed and α varying. Figure 10 shows contour maps over k_0 and α for metrics based on each of the first four harmonics. These indicate that $\alpha = 0.50$ is indeed a reasonable assumption.

5. Summary and future work

The work discussed in this paper has two major outcomes, one concerning the science and the other the e-Science tools. In terms of the science, a methodology has been developed that has produced new understanding for the $\text{Fc}^{0/+}$ process

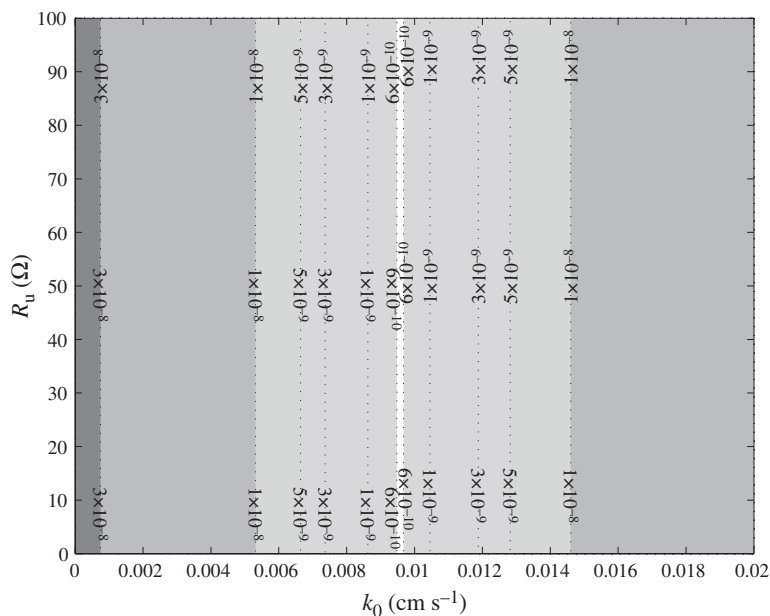


Figure 9. Contour map for the metric E3 against k_0 and R_u for potassium ferricyanide.

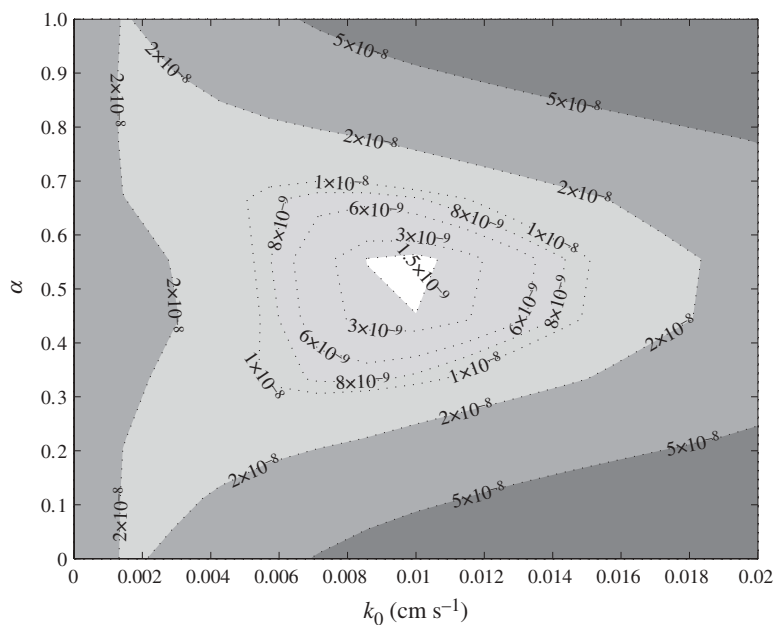


Figure 10. Contour map for the metric E3 against k_0 and α for reduction of ferricyanide.

and validates literature data for the $[\text{Fe}(\text{CN}_6)]^{3-/4-}$ process, which promises widespread applicability. As a result, we plan further voltammetric experiments. For example, we will experiment with square, triangular and sawtooth waveforms, and multiple sine waves of different amplitudes and frequencies [3]. Further, we

will introduce more complex electrochemical mechanisms, where chemical steps are coupled to electron transfer, and schemes with surface-confined reactions, where heterogeneity of the electrode surface allows for a range of k_0 , E° and α values [21]. These outcomes would have benefited from a multi-disciplinary approach to the work (bringing chemists and computer scientists together), which demonstrates the significant achievements attainable by e-Science in general.

In terms of the e-Science tools, we have (again) demonstrated the power of the Nimrod tool family and have illustrated how the tools used are sufficiently flexible to quickly modify the experiments as needed. However, more modifications are suggested. We would like to quantify the sensitivity of the various harmonics on the parameters. Other optimization algorithms in the Nimrod/O suite may be trialled. Three-dimensional contour maps (or at least their projections onto two) may reveal further dependence between the parameters. In all of these studies, the availability of e-Science tools implies that the expertise required will be chemical and mathematical, rather than computer science.

This work used just a single eight-core compute node, sufficient for these purposes, as no experiment lasted more than an hour. But, under Nimrod, it is a simple task to extend the testbed to use much larger clusters, grids and clouds. This would allow for multiple concurrent experiments and speed up the sweep operation in particular. Another possible development would be to bundle the parameter fitting software with the data acquisition software and to run both on one multi-core machine. The experimenter could then run the experiment and obtain parameter estimates in one operation.

One of the more general observations of our work is the iterative nature of an e-Science project, involving dialogue between scientists from different disciplines. Our initial thoughts on what would work were clearly naive and were refined throughout the experiment. More importantly, observations made by different members of the multi-disciplinary team were often fed back into the design of the next experiment.

Appendix. Experimental set-up

This appendix is provided for those interested in the details of the chemistry and is separated from the rest of the paper, which is focused on the e-Science infrastructure.

A standard three-electrode cell was employed in all electrochemical measurements. Glassy carbon disc electrodes (Bioanalytical Systems) with 3 mm diameter and platinum wire were employed as working and counter-electrodes, respectively. Prior to use in voltammetric experiments, the glassy carbon electrode was polished with an aqueous 0.3 μm alumina suspension on a polishing cloth (Buehler), rinsed with water, sonicated to remove excess alumina, before a final rinse with water and drying in nitrogen. For experiments in acetonitrile (0.1 M Bu_4NPF_6) solvent, a silver wire quasi-reference electrode was used. For aqueous solution, an $\text{Ag}/\text{AgCl}/\text{NaCl}$ (3 M) reference electrode was employed.

The large-amplitude FT AC cyclic voltammetric experiments were performed using instrumentation described in more detail elsewhere [17]. The instrumentation diagram is presented in figure 11. Analogue signals are produced with a stereo 18-bit digital-to-analogue converter (DAC) running at 39 kHz.

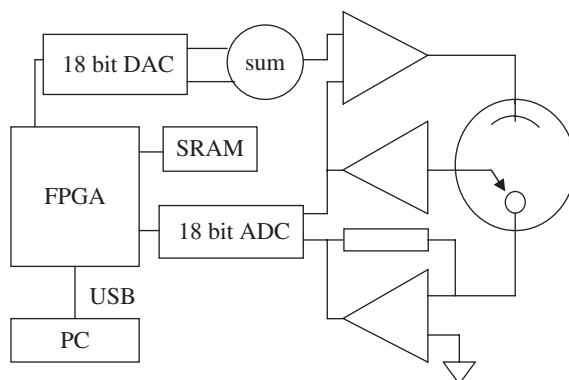


Figure 11. Data acquisition for Fourier voltammetry.

The resulting voltage and current are digitized with a stereo 18-bit analogue-to-digital converter (ADC). One channel of the DAC produces a ± 3 V ramp that is electrically summed with a ± 300 mV perturbation signal produced by the other channel. This combined signal is applied to a conventional analogue potentiostat configuration. The reference electrode is buffered before being used within the analogue feedback loop. This voltage is also fed to one of the channels of the ADC, which gives a measure of the actual cell voltage. The cell current is converted to a voltage using an appropriately configured operational amplifier, which is then connected to the other channel of the ADC. The synchronous control of the DAC and ADC is achieved using a field programmable gate array (FPGA). The FPGA is also connected to an external memory device (static RAM) to provide buffering of both the applied and measured signals. This is necessary as the Universal Serial Bus (USB) connection to the personal computer provides and accepts data in bursts.

The experimental data for the ferrocene AC test were obtained with the FT voltammetric instrumentation using a sine wave perturbation with a frequency of 9.02 Hz and amplitude of 80 mV, and the number of data points are 16 384, ramp potential -0.3 V to $+1.0$ V and back to -0.3 V, scan rate 0.9686 s^{-1} , at temperature 295 K. Experimental conditions for ferricyanide employed the same frequency, amplitude and data points as in the case of ferrocene, with the ramp potential $+0.6$ V to -0.1 V and back to $+0.6$ V, scan rate 0.05215 s^{-1} at 293 K.

References

- 1 Peachey, T. C. 2009 A sample of eScience applications using MeSSAGE lab tools. See <http://messagelab.monash.edu.au/EScienceApplications> (accessed December 2010).
- 2 Bard, A. J. & Faulkner, L. R. 1980 *Electrochemical methods: fundamentals and applications*, New York, NY: Wiley.
- 3 Bond, A. M., Duffy, N. W., Guo, S.-X., Zhang, J. & Elton, D. 2005 Changing the look of voltammetry. *Anal. Chem.* **77**, 186A–195A. (doi:10.1021/ac053370k)
- 4 Box, G. E. 1994 Statistics and quality improvement. *J. R. Statist. Soc. A* **157**, 209–229. (doi:10.2307/2983359)

- 5 Sher, A. A., Bond A. M., Gavaghan, D. J., Harriman, K., Feldberg, S. W., Duffy, N. W., Guo, S.-X. & Zhang, J. 2004 Resistance, capacitance, and electrode kinetic effects in Fourier-transformed large-amplitude sinusoidal voltammetry: emergence of powerful and intuitively obvious tools for recognition of patterns of behavior. *Anal. Chem.* **76**, 6214–6228. (doi:10.1021/ac0495337)
- 6 Sher, A. A., Bond A. M., Gavaghan, D. J., Gillow, K., Duffy, N. W., Guo, S.-X. & Zhang, J. 2005 Fourier transformed large amplitude square-wave voltammetry as an alternative to impedance spectroscopy: evaluation of resistance, capacitance and electrode kinetic effects via an heuristic approach. *Electroanalysis* **17**, 1450–1462. (doi:10.1002/elan.200403293)
- 7 Abramson, D., Soscic, R., Giddy, J. & Hall, B. 1995 Nimrod: a tool for performing parametrised simulations using distributed workstations. In *The 4th IEEE Symp. on High Performance Distributed Computing, Virginia, August 1995*, pp. 112–121.
- 8 Abramson, D., Giddy, J. & Kotler, L. 2000 High performance parametric modeling with Nimrod/G: killer application for the global grid? In *Int. Parallel and Distributed Processing Symp. (IPDPS), Cancun, Mexico, May 2000*, pp. 520–528. .
- 9 Bethwaite, B., Abramson, D., Bohnert, F., Garic, S., Enticott, C. & Peachey, T. Mixing grids and clouds: high-throughput science using the Nimrod tool family. In *Cloud computing: principles, systems and applications*. Berlin, Germany: Springer.
- 10 Ravindran, A., Ragsdell, K. M. & Reklaitis, G. V. 2006 *Engineering optimization: methods and applications*. New York, NY: Wiley.
- 11 Abramson, D., Peachey, T. & Lewis, A. 2006 Model optimization and parameter estimation with Nimrod/O. In *The Int. Conf. on Computational Science, University of Reading, May 2006*.
- 12 Yu, J. & Buyya, R. 2005 A taxonomy of workflow management systems for grid computing. *J. Grid Comput.* **3**, 171–200. (doi:10.1007/s10723-005-9010-8)
- 13 Ludäscher, B., Altintas, I., Berkley, C., Higgins, D., Jaeger, E., Jones, M., Lee, E. A., Tao, J. & Zhao, Y. 2006 Scientific workflow management and the Kepler system. *Concurr. Comput.: Pract. Exper.* **18**, 1039–1065 (Special Issue: Workflow in Grid Systems). (doi:10.1002/cpe.994)
- 14 Abramson, D., Enticott, C. & Altintas, I. 2008 Nimrod/K: towards massively parallel dynamic grid workflows. In *IEEE SuperComputing 2008, Austin, TX*.
- 15 Abramson, D., Bethwaite, B., Enticott, C., Garic, S., Peachey, T., Michailova, A. & Amirriazi, S. 2010 Embedding optimization in computational science workflows. *J. Comput. Sci.* **1**, 41–47. (doi:10.1016/j.jocs.2010.04.002)
- 16 Nelder, J. A. & Mead, R. 1965 A simplex method for function minimization. *Comput. J.* **7**, 308–313.
- 17 Bond, A. M., Duffy, N. W., Elton, D. M. & Fleming, B. D. 2009 Characterization of nonlinear background components in voltammetry by use of large amplitude periodic perturbations and Fourier transform analysis. *Anal. Chem.* **81**, 8801–8808. (doi:10.1021/ac901318r)
- 18 Tsierkezos, N. G. & Ritter, U. 2010 Electrochemical impedance spectroscopy and cyclic voltammetry of ferrocene in acetonitrile/acetone system. *J. Appl. Electrochem.* **40**, 409–417. (doi:10.1007/s10800-009-0011-3)
- 19 Xiao, L., Dickinson, E. J. F., Wildgoose, G. G. & Compton, R. G. 2010 A comparison of electron transfer kinetics of three common carbon electrode surfaces in acetonitrile and in room temperature ionic liquid 1-butyl-3-methylimidazolium hexafluorophosphate: correlation to surface structure and the limit of the diffusion domain approximation. *Electroanalysis* **22**, 269–276. (doi:10.1002/elan.200900399)
- 20 Hu, I. F., Karweik, D. H. & Kuwana, T. 1985 Activation and deactivation of glassy carbon electrodes. *J. Electroanal. Chem.* **188**, 59. (doi:10.1016/S0022-0728(85)80050-4)
- 21 Lee, C.-Y., Bullock, J. P., Kennedy, G. F. & Bond, A. M. 2010 Effects of coupled homogeneous chemical reactions on the response of large-amplitude AC voltammetry: extraction of kinetic and mechanistic information by Fourier transform analysis of higher harmonic data. *J. Phys. Chem. A* **114**, 10122–10134. (doi:10.1021/jp105626z)