Faraday Discussions

Cite this: *Faraday Discuss.*, 2022, **233**, 58

DISCUSSIONS

View Article Online
View Journal | View Issue

Electrochemical data mining: from information to knowledge: general discussion

Tim Albrecht, Xiangkun Elvis Cao, Dongfei Chen, Manuel Corva, Martin A. Edwards, Andrew Ewing, Stefano Fornasaro, J. Justin Gooding, Luke Gundry, Ayumi Hirano-Iwata, Grant Jeffcoat, Ali Reza Kamali, Frédéric Kanoufi, Serge G. Lemay, Ndrina Limani, Steven Linfield, Xu Liu, Si-Min Lu, Gabriel N. Meloni, Zhongqun Tian, Kristina Tschulik, Swathi Naidu Vakamulla Raghu, Hui Wei and Yi-Lun Ying

DOI: 10.1039/d2fd90001g

Yi-Lun Ying opened discussion of the introductory lecture by Justin Gooding: The nanointerface needs more comprehensive models to describe the interaction networks among water and ions. Would you please comment on how these models and understanding guide the applications of nanoelectrochemistry?

Justin Gooding answered: This is a very difficult question which I don't think we know enough yet to give you a clear answer. What we do know is that with surfaces inside the substrate channels of our enzyme-like nanoparticles we have two to three layers of ice like water. Inside the channels that might mean that all the water is in an ice-like state. How this affects transport in the channel and the electrical double layer in electrochemical systems are big questions for us. Certainly, regarding transport, there is evidence for proton migration inside these channels and possibly sodium and potassium if they are present. But in our systems our materials are not well defined enough to tease out such effects. This was what my point was about regarding needing better defined materials to really understand the nature of the nanointerface in nanoparticle systems. There is however a lot of information we can extract from studies performed in the transport of species through nanochannels. I don't think we yet have an understanding though for case that I described where the nanochannels themselves are also reactive.[1,2]

[1] T. M. Benedetti, C. Andronescu, S. Cheong, P. Wilde, J. Wordsworth, M. Kientz, R. D. Tilley, W. Schuhmann and J. J. Gooding, Electrocatalytic nanoparticles that mimic the three dimensional geometric architecture of enzymes: nanozymes, *J. Am. Chem. Soc.*, 2018, **140**, 13449–13455, DOI: 10.1021/jacs.8b08664.

2 J. Wordsworth, T. M. Benedetti, A. Alinezhad, R. D. Tilley, M. A. Edwards, W. Schuhmann and J. J. Gooding, The importance of nanoscale confinement to electrocatalytic performance, *Chem. Sci.*, 2020, **11**, 1233–1240, DOI: 10.1039/c9sc05611d.

Si-Min Lu enquired: I have a question on the COMSOL modeling in your study. You mentioned that the migration plays a important role in the electrocatalysis. I am wondering if there is any possibility that the electric convection could enhance the mass transport in the confined channel as the electroosmotic flow may exist in the channel where the double layer is large enough. The clever marriage between finite element modeling and molecular dynamics (MD) simulation may be a good way for unveiling the double layer structure in the nanoconfinement.

Justin Gooding replied: This is another very good question that I don't think we have a definitive answer to yet. What we believe our COMSOL modelling shows is there is some active transport of protons into the channels but also there is a possibility that the water structure inside the channels is influencing proton transport. All we can say with the COMSOL modelling, as it is a continuum model, is the increase in catalytic activity we see is consistent with the active transport of protons going into the channel. In our model however, it approximates the surface potential as charge on the channel walls and the diffuse double layer is composed of point charges. So we do need molecular dynamics simulations inside the channels to really understand what the relative effects might be. So I agree that a combination of COMSOL modelling with MD simulation would be a good way to start to understand such systems, although we also need better defined experimental systems and we do need MD models to include the consumption of the protons within the channel.

Swathi Naidu Vakamulla Raghu said: It was mentioned that in array-like structures each channel will have its own micro-environment of sorts. Is this the case say for a homogeneous structure such as with metal oxide nanotubes?

Justin Gooding responded: So why I make this statement is from our work on porous silicon,[1] so not from electrochemical systems but optical devices. In the case of porous silicon you have well-defined columnar pores, so each nanopore is an isolated pore. There what we see is the materials act like a sponge, if we have low concentrations of analyte outside the structure, because each nanopore is a separate diffusional system, we get high concentrations of analyte inside the nanoporous structure. This is because each open channel is in equilibrium with the exterior solution so if a molecule goes into the channel, the other channels nearby do not feel the effect of that molecule entering the structure. However, in a mesoporous structure where each pore is connected with the others, when a molecule enters one pore this is felt by the other pores as they are all interconnected within the nanostructure.[2]

1 K. A. Kilian, T. Böcking, K. Gaus, J. King-Lacroix, M. Gal and J. J. Gooding, Hybrid lipid bilayers in nanostructured silicon: a biomimetic mesoporous scaffold for optical detection of cholera toxin, *Chem. Commun.*, 2007, 1936–1938, DOI: 10.1039/b702762a.

2 J. Wordsworth, T. M. Benedetti, A. Alinezhad, R. D. Tilley, M. A. Edwards, W. Schuhmann and J. J. Gooding, The importance of nanoscale confinement to electrocatalytic performance, *Chem. Sci.*, 2020, **11**, 1233–1240, DOI: 10.1039/c9sc05611d.

Hui Wei asked: You showed nice examples of how to design enzymatic channels in nanozymes. I am wondering whether you have considered the design of more flexible (softer) channels in (inorganic) nanomaterials? We know natural enzymes are usually softer and more flexible.

Another question is: you demonstrated the nanoconfinement effect in electrochemical catalysis systems. Would the thermally driven catalysis have similar effects as shown in electrochemistry driven catalysis?

My third question is: By using external magnetic force and magnetic particles, you can improve the detection sensitivity. For the analytes of ultralow concentration, how would one enhance the interaction between analytes and the antibodies (or other recognition elements) on a magnetic particle?

Justin Gooding replied: With the rigid structures that we have made one could use the small size of the channels to enhance selectivity *via* size exclusion. The idea of flexible channels to mimic allosteric enzymes is not something we have considered yet. What we would like to do is make nanoparticles where the reaction only occurs at the end of the channels, rather than also reacting along the channel walls, to better mimic enzymes. We feel this would deconvolute some of the uncertainties related to what is occurring in our system. We don't currently know how to make the materials flexible anyway. We are actually not trying to mimic every aspect of an enzyme but rather with that work we are trying to just mimic how an enzyme controls the solution properties where the reaction proceeds to facilitate that reaction. That is what we would like to know how to do because we think that could give us significant advantages in electrocatalysis. Regarding the second question, in many ways I think similar effects can occur in non-electrochemically driven systems, but in our system we can use potential to control the reaction which is an option other approaches may not provide to the same extent. Regarding the third question and facilitating interactions between the analyte and antibodies in magnetic nanoparticles. We still rely on the affinity between the antibody and the analyte but by having many antibodies on the nanoparticle we actually enhance the chances of an analyte to remain bound to the nanoparticle as we have discussed previously.[1]

1 Y. F. Wu, D. Bennett, R. D. Tilley and J. J. Gooding, How nanoparticles transform single molecule measurements into quantitative sensors, *Adv. Mater.*, 2020, **32**, 1904339.

Ndrina Limani enquired: In the ‘Electrocatalysis’ part of your presentation (DOI: 10.1039/d1fd00088h), some ORR and OER electrocatalytic activities were presented, mainly to illustrate the performance of novel nanomaterials. However, there are several methods that could be used to assess the electrocatalytic activity (SECM, RRDE…) and I was wondering how was it assessed in those cases?

Justin Gooding answered: To determine the electrocatalytic activity we usually used rotating disc electrodes but of course also had to determine the electroactive area. Most frequently for Pt surfaces we used Cu underpotential deposition (UPD). The reasoning for using Cu UPD rather than H UPD was for two different reasons. The first reason is that because of the high curvature on some of the nanoparticle systems we deal with there is a possibility of strain effects that will influence the H UPD as we have shown in ref 1. The second reason is for the nanozyme-style particles that mimic the 3D geometry of enzymes the hydrogen can also pass through the surfactant layer around the nanoparticles, but the copper does not, so we get the area of the substrate channel only. This is discussed in ref. 2.

1 A. Alinezhad, L. Gloag, T. M. Benedetti, S. Cheong, R. F. Webster, M. Roelsgaard, B. B. Iversen, W. Schuhmann, J. J. Gooding and R. D. Tilley, Direct growth of highly strained Pt islands on branched Ni nanoparticles for improved hydrogen evolution reaction activity, *J. Am. Chem. Soc.*, 2019, **141**, 16202–16207.
2 T. M. Benedetti, C. Andronescu, S. Cheong, P. Wilde, J. Wordsworth, M. Kientz, R. D. Tilley, W. Schuhmann and J. J. Gooding, Electrocatalytic nanoparticles that mimic the three-dimensional geometric architecture of enzymes: nanozymes, *J. Am. Chem. Soc.*, 2018, **140**, 13449–13455.

Frédéric Kanoufi communicated: This was a very inspiring talk showing a wide spectrum of single entity inspection and the challenges to face. The inspection of nanoporous nanoparticle activity is really inspiring. You particularly mentioned the effect of migration within single nanoparticle nanopores. It is a major field of research in water and electrolyte transport in single nanostructures (*e.g.* nanotubes, see ref. 1). In this field the many works suggest osmotic transport of water within nanoconfined regions. Do you think this might happen in your system (it would increase the electrocatalysis mass transport)? It is definitely an interesting aspect to unveil by electrochemistry or to combine with electrocatalysis.

1 K. Falk, F. Sedlmeier, L. Joly, R. R. Netz and L. Bocquet, *Nano Lett.*, 2010, **10**, 4067.

Justin Gooding communicated in reply: Yes I agree that osmotic transport could be occurring in our nanoparticles with nanopores and I mentioned studies in nanochannels and nanotubes. When I was referring to migration, I think it is more accurate for us to say that our results are consistent with active transport of protons into the channel without us having the control over the experimental system to tease out which of the various possibilities of that active transport could be the dominant effect.

Ali Reza Kamali opened discussion of the paper by Yi-Lun Ying: Very interesting work! I am wondering whether you have studied the performance of the device in non-aqueous phases too?

Yi-Lun Ying answered: We have not studied the performance of our amplifier system in non-aqueous phases yet. Our amplifier array could be used to work in other conditions such as organic phase and solid phase if it is adopted to a suitable electrode array.

Si-Min Lu enquired: The instrumental advances are pivotal to single entity electrochemistry at the nanoscale as the electrochemical instrumentation with high bandwidth can resolve more accurate information on mass transport and electron transfer of single entities at the confined interface. This study enables parallel nanoelectrochemical measurements. I am interested in the chip architecture, what is the distances between each electrode in the array? Why? Have you optimized the distance by simulation?

Yi-Lun Ying replied: We are currently working on optimizing the distance between electrodes. The interference between electrodes would be diminished when the distance is set to >0.5 mm according to previous work.[1] Here, the distance set into 1.27 mm which was accessible to the fabrication of the array chip.

1 M. Akhtari, D. Emin, B. M. Ellingson, D. Woodworth, A. Frew and G. W. Mathern, *J. Appl. Phys.*, 2016, **119**, 064701.

Zhongqun Tian asked: In the electrochemical environment, a single entity is not isolated and its behavior may be influenced by many factors. So its characteristic current including noise may provide some information about the nanopore/nanoelectrode such as edge, shape, soft and hard electrode materials, dynamical change of modified layer and biomolecules, *etc.* The noise analysis may be important for deep understanding of the mechanism and getting a full physico-chemical picture of the single entity. So I wonder how much useful information on the sub-entity level and/or molecular/atomic level one can extract from the so called noise? And how to technically distinguish the real useful noise from thermal and artificial noises generated by electronics and instrument? If possible could you summarize or predict the methods and approaches of the noise analysis in this field?

Yi-Lun Ying responded: Indeed, much information could be extracted from the noise on the sub-entity level and/or molecular/atomic level. The algorithm for big data analysis is useful to calculate dynamic physico-chemical properties from signals. We summarized our developed algorithm into the following aspects:

(i) Locating and recovering the attenuated signal. The duration time of the filter distorts the blockages. This attenuation could be calibrated by using a second-order-differential-based calibration. Then, the current amplitude of the blockages was recovered with an integrated method.[1,2] The results showed that the relative error for evaluating the blockages with a duration as short as 0.01 ms was significantly decreased to <20% which was lower than the conventional method (>800%). In order to automatically locate the short blockages from the highly noisy data, we developed a Modified Hidden Markov Model (MHMM) which includes the Fuzzy c-Means to initialize the HMM parameters and Viterbi training algorithm to optimize the HMM. This algorithm shows strong tolerance to the noise which enables detection of nanopore blockages at the highest bandwidth of the amplifier.[3]

(ii) Identification of similar blockades from the mixture. The analytes that differ with only a subgroup (*e.g.* methylcytosine and cytosine) often exhibit similar blockage currents or/and durations. To improve the sensing ability of nanopores, a shapelet-based machine learning approach was used to discriminate mixed analytes that exhibit nearly identical blockage current amplitude and durations.[4] This algorithm would facilitate the application of nanopore techniques in complex real samples. Moreover, the adaBoost model improved the capacity of the nanopore in efficient identification of single nucleotide differences.[5]

(iii) Tracing time-series single-molecule characteristics. The transition conformation of a single molecule usually generates time-series spike-like current fluctuations which could be described by using a four-stage Markov chain model.[6] This method built discrete-state stochastic models for describing long-term statistical dynamics. Therefore, the combination of nanopores with the Markov chain model clarified the transition pathways of the β-hairpin peptide. For single nanoparticles collision experiments, the envelope algorithm was developed to identify multi-peaks and gave the integrated charges calculation.[7]

(iv) Mining ion-interaction networks from the "noise". The transport of single analyte and ions through a biological nanopore is governed by interaction networks among the confined ions, analyte and residues at the inner surface of the nanopore. However, the ion interaction networks only generate very weak ionic fluctuations which are usually buried in the blockage noise. In order to achieve qualitative and spectral analysis of such weak fluctuations, we developed an advanced frequency method based on the Hilbert–Huang transform.[8] This method quantitatively bridges the gap between frequency characteristics and ion interaction networks inside the nanopore. Especially, the peak frequency (f_m) and its amplitude (a_m) were used as a frequency fingerprint to describe and predict dynamic equilibriums for ion interaction networks. We anticipate that the frequency characteristics could be extended to study interaction networks for various analytes, allowing the nanopore-based biopolymer sequencing.

1 Z. Gu, Y.-L. Ying, C. Cao, P. He and Y.-T. Long, *Anal. Chem.*, 2015, **87**, 907.
2 Z. Gu, Y.-L. Ying, C. Cao, P. He and Y.-T. Long, *Anal. Chem.*, 2015, **87**, 10653.
3 J. Zhang, X. Liu, Y.-L. Ying, Z. Gu, F.-N. Meng and Y.-T. Long, *Nanoscale*, 2017, **9**, 3458, DOI: 10.1039/c6nr09135k.
4 Z.-X. Wei, Y.-L. Ying, M.-Y. Li, J. Yang, J.-L. Zhou, H.-F. Wang, B.-Y. Yan and Y.-T. Long, *Anal. Chem.*, 2019, **91**, 10033.
5 X.-J. Sui, M.-Y. Li, Y.-L. Ying, B.-Y. Yan, H.-F. Wang, J.-L. Zhou, Z. Gu and Y.-T. Long, *J. Anal. Test.*, 2019, **3**, 134.
6 S.-C. Liu, Y.-L. Ying, W.-H. Li, Y.-J. Wan and Y.-T. Long, *Chem. Sci.*, 2021, **12**, 3282, DOI: 10.1039/d0sc06106a.
7 C.-Y. Xu, R.-J. Yu, X. Ni, S.-W. Xu, X.-X. Fu, Y.-J. Wan, Y.-L. Ying and Y.-T. Long, *Chin. J. Chem.*, 2021, **39**, 1936.
8 X. Li, Y.-L. Ying, X.-X. Fu, Y.-J. Wan and Y.-T. Long, *Angew. Chem., Int. Ed.*, 2021, **60**, 24582.

Ayumi Hirano-Iwata commented: The current noise level is also dependent on the sampling frequency of recordings. If you use a higher sampling frequency, the noise will also increase. Because the sampling frequency is dependent on how fast the target reactions are, it would be helpful if you could show us the time resolution and noise level that are required for your target reactions.

Yi-Lun Ying responded: The noise of the chemical system and circuit system is mainly affected by three factors of the instrument: bandwidth of amplifiers, cut-off frequency of the low-pass filter and the sampling frequency. Moreover, the capacitance from the electrode (C_E) and connection wires (C_W) also contribute to the noise. To achieve a satisfactory signal-to-noise ratio (SNR) in the electrochemical measurement, a low-pass filter was adopted, which gave the cut-off frequency of 5 kHz. The noise level of the single-entity electrochemical measurement in our system at 5 kHz filter was about 4–5 pA RMS. The time resolution was about 0.2 ms. In order to avoid the signal distortion, the value of the sampling rate usually should be set to 2.56 times higher than that of the cut-off frequency of the filter according to the Nyquist sampling rate. Therefore, the sampling rate in our system should be set above 12.8 kHz. Here, we set the sampling rate to as high as 100 kHz which is sufficient enough to avoid the signal distortion from analog-to-digital conversion.

Martin A. Edwards asked: In your system, do you know how much of the noise is coming from the amplifiers *vs.* the electrodes and chemistry of the system? You present results from relatively large microelectrodes (200 μm). I expect the noise from the electrodes and their capacitance to scale with their area (radius squared) – thus there might be great opportunities to drastically lower the capacitance/noise. Do you anticipate that this would be useful?

Yi-Lun Ying answered: That would be useful. In our system, the noise mainly originates from the electrodes and chemistry of system. The electrode interface capacitance (C_E) mainly contributes to the noise. The capacitance is proportional to the area of the electrode. Therefore, the reduction of the area of the electrode will be very useful to lower the noise.

Andrew Ewing addressed Martin A. Edwards and Yi-Lun Ying: Further to the previous question by Martin A. Edwards: that implies interference noise, right? Need a noise spectrum analysis? 1 pA is good for a 200 μm electrode, but not for a smaller electrode.

Yi-Lun Ying replied: For the experiment, we should compromise between the bandwidth in the whole recording system and the required signal-to-noise ratio.

Martin A. Edwards replied: Yes, noise spectrum is what I was thinking. Interference (background processes) was what I was imagining, but there are likely also electrical components due to the double-layer capacitance that might have some impact. Yes, 1 pA seems good for a 200 μm electrode, although as you are well aware, we should also consider the frequency we are assessing it at to make truly meaningful comparisons.

Serge Lemay enquired: The compensation stage also amplifies noise from the first amplifier stage in the additional frequency window, correct? Could you please quantify the impact that this has on the overall performance of the circuit?

Yi-Lun Ying answered: The compensation circuit indeed amplifies the input noise and noise from the first amplifier stage. But the compensation circuit only amplifies the input signal at the high frequency region of >1 kHz, see Fig. 1 here. In this region, the noise is from dielectric loss and capacitance (Fig. 1b in the paper; DOI: 10.1039/d1fd00055a). In other words, the compensation circuit compensates the gain loss of the first-stage amplifier at the high frequency region. As shown in Fig. 3c in the paper (DOI: 10.1039/d1fd00055a), the −3 db bandwidth was increased from 1 kHz to over 100 kHz. In order to further suppress the noise, the low-pass filter could be used. For example, in our single-molecule detection assay, a 5 kHz low-pass filter was employed to further reduce the noise in the high frequency region.

Andrew Ewing asked: Have you analysed the noise types using a frequency spectrum analyser? Just off the top of the head, the success of the 2-stage amplifier I think suggests you are limited by shot noise, but 1 pA is not really to that limit. What does the frequency spectrum look like if you put in high frequencies?

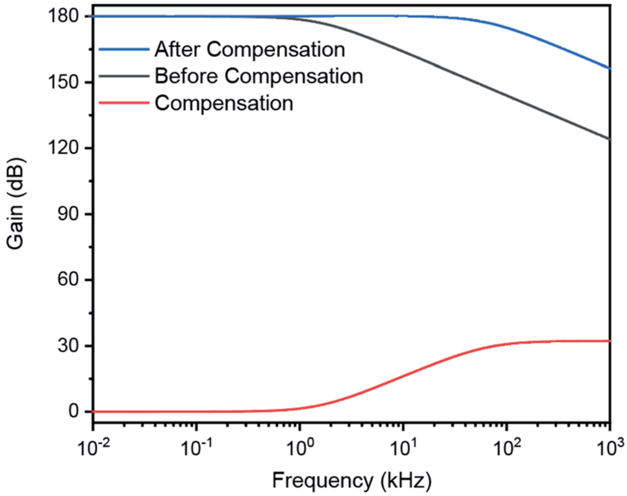

Fig. 1 Bandwidth simulation for the amplifier circuit performance. Bandwidth of four-channel amplifier circuits without compensation (black), the compensation circuit (red) and after compensation (blue).

Yi-Lun Ying responded: The noise in the experiments depends not only on the amplifier but also on the electrode array and the connection wire. Currently, we adopt an electrode array with a relatively large electrode of diameter of 150 μm (Fig. 2b in the paper; DOI: 10.1039/d1fd00055a). This electrode array caused additional noise in the measurement. Therefore, we adopted a 5 kHz Bessel low-pass filter to remove the noise at high frequencies. The further incorporation of a nanoelectrode array would be beneficial for the noise reduction.

Kristina Tschulik said: Thank you for sharing your work on an amplifier for high bandwidth, low current applications. Developments in this direction will be essential to overcome existing limitations in single entity electrochemistry. Could you please specify what is the maximum filter frequency you have been able to use for filtering in nanoparticle impact studies to get appreciable signal-to-noise ratios for the example systems you present in your paper (DOI: 10.1039/d1fd00055a)?

Yi-Lun Ying replied: This is a really good question. We are exploring the potential highest bandwidth for recording single-nanoparticle collisions at the electrode. The whole bandwidth of the single-nanoparticle collisions depends not only on the amplifier but also on the electrode. The double layer capacitance of the electrode is proportional to the electrode area. Consider the electrode capacitance (C_E) as a simple parallel plate capacitor given by

$$C_E = \frac{\varepsilon A}{d}$$

where A is the area of the electrode and d is the thickness of the double layer. Therefore, the greatly reduced surface area of the micro-/nanoelectrode results in a small C. According to our previous study, at the same experimental conditions, the root mean square (RMS) noise decreased to 2.7 pA at 5 kHz filter as the diameter of the electrode reduced to 480 nm.[1] Moreover, a 30 nm close-type wireless nanopore electrode achieved a RMS of 1.3 pA in AgNPs collision experiments.[2] Usually, the signal-to-noise ratio (SNR) should exceed 3 for the following identification of each signal from the current recording. To achieve a satisfactory SNR, the bandwidth of the whole system is expected to reach >5 kHz if a nanoelectrode with low capacitance is adopted. Note that the defined maximum bandwidth ($B_{\max}$) at which the SNR exceeds a minimum value ($\mathrm{SNR}_{\min}$) is given by

$$B_{\max} \approx \left(\frac{\Delta I \sqrt{3}}{\mathrm{SNR}_{\min} \times 2\pi (C_\mathrm{E} + C_\mathrm{W} + C_\mathrm{In} + C_\mathrm{F}) v_\mathrm{n}} \right)^{2/3}$$

where C_W is capacitance contributions from the electrode connecting wire, C_In is the amplifier input and C_F is the amplifier parasitic capacitance (C_F) of a feedback resistor (R_F). If C_E and C_W could be reduced to 1 pF together with the low C_F of this designed amplifier (0.1 pF), it may be possible for the bandwidth to reach a maximum value of 100 kHz. However, such a high bandwidth needs continuous efforts in amplifier design, chip connection and electrode array manufacturing. Also, the mass transport of the redox molecule should be considered when we perform the high bandwidth recording.

1 Y. Zheyao, M. Wei, Y. Yilun and L. Yitao, *Acta Chim. Sin.*, 2017, **75**, 671.
2 R.-J. Yu, S.-W. Xu, S. Paul, Y.-L. Ying, L.-F. Cui, H. Daiguji, W.-L. Hsu and Y.-T. Long, *ACS Sens.*, 2021, **6**, 335.

Kristina Tschulik noted: For readers who want to test the bandwidth of their lab equipment (which is not always transparent *e.g.* for commercial potentiostats), I would like to suggest a hands-on guide for electrochemists we've developed.[1]

1 K. Kanokkanchana, E. N. Saw and K. Tschulik, *ChemElectroChem*, 2018, **5**, 3000–3005, DOI: 10.1002/celc.201800738.

Yi-Lun Ying added: The temporal resolution of the single-entity electrochemical measurement is inversely related to the bandwidth in the recording system but also is limited by the required signal-to-noise ratio (SNR). Various research studies have discussed the impact of bandwidth, including but not limited to ref. 1 and 2. Our group has also elaborated the bandwidth impact in perspective and review papers.[3–5]

1 E. Kätelhön, E. E. L. Tanner, C. Batchelor-McAuley and R. G. Compton, Destructive nanoimpacts: What information can be extracted from spike shapes?, *Electrochim. Acta*, 2016, **199**, 297.
2 R. Gao, M. A. Edwards, J. M. Harris and H. S. White, Shot noise sets the limit of quantification in electrochemical measurements, *Curr. Opin. Electrochem.*, 2020, **22**, 170.
3 J. Wang, Y.-L. Ying, C.-B. Zhong, L.-M. Zhang, F. Yan and Y.-T. Long, Instrumentational implementation for parallelized nanopore electrochemical measurements, *Analyst*, 2021, **146**, 4111, DOI: 10.1039/d1an00471a.
4 Z. Gu, Y.-L. Ying and Y.-T. Long, Nanopore sensing system for high-throughput single molecular analysis, *Sci. China: Chem.*, 2018, **61**, 1483.
5 Y.-L. Ying, Z. Ding, D. Zhan and Y.-T. Long, Advanced electroanalytical chemistry at nanoelectrodes, *Chem. Sci.*, 2017, **8**, 3338, DOI: 10.1039/c7sc00433h.

Gabriel Meloni questioned: In Fig. 4b of the paper (DOI: 10.1039/d1fd00055a) it looks like there is still some distortion in the transient (looks a bit like a discharge) while you are still away from the frequency limitation of your amplifier. Is that due to an under compensation of the compensating circuit? And how do you set that compensation? Also, in the simulations you used a top-end parasitic capacitance of 0.5 pF, isn't that too small for a real system, considering the length of the wires you have in Fig. 2 of the paper? Final point, any comments on power supply stability? For me not many people think about it.

Yi-Lun Ying replied: (i) About the signal distortion. In order to achieve a high SNR, we set the low-pass filter at 5 kHz. Therefore, the maximum bandwidth of the whole measurement system is 5 kHz. The possible distortion originates from the low-pass filter. However, the current shape of the recorded signal is similar to the previous single-nanoparticle collision for H_2 evolution reactions using a commercial amplifier at 5 kHz.[1] Therefore, each channel of our amplifier array showed comparable performance to that from a commercially available single-channel amplifier (*e.g.* Axon 200B).

(ii) About compensation. By adjusting the parameters of R_{com}, C_{com} and R_{adj} in the compensation circuit, the −3 dB bandwidth of the amplifier could reach as high as over 100 kHz (Fig. 3c of the paper; DOI: 10.1039/d1fd00055a). The final values of R_{com}, C_{com}, R_{adj} were set to 20 Ω, 50 nF and 150 Ω, respectively.

(iii) About parasitic capacitance. Fig. 3d in the paper simulates the bandwidth reduction from crosstalk of the multi-channels and the structure design of the board. It is reasonable to set the parasitic capacitance of feedback resistance to 0.5 pF based on our circuit board design. This value is not the total capacitance for the whole measurement system in single-particle collisions.

(iv) About power supply stability. The power supply was stabilized by a low dropout regulator (LDO) which is an essential power management circuit. The voltage noise is suppressed to <20 μV RMS.

[1] M. Chen, S.-M. Lu, Y.-Y. Peng, Z. Ding and Y.-T. Long, *Chem.–Eur. J.*, 2021, **27**, 11799.

Manuel Corva asked: I am interested in the Pt nanoimpacts you present in the paper (DOI: 10.1039/d1fd00055a). Why is no HER steady-state current observed despite working in 10 mM $HClO_4$?

Yi-Lun Ying responded: The rapid current spike is due to the initial hydrogen adsorption process (the Volmer step, $H^+ + e^- \rightarrow H_{\mathrm{ads}}$), while the slow decay is due to the hydrogen evolution process. At potentials ranging from 0 to −400 mV, the overpotential is insufficient to drive the evolution reaction, and therefore the hydrogen adsorption will proceed until the entire Pt surface is covered by a monolayer of H_{ads} (Fig. 4B of the paper; DOI: 10.1039/d1fd00055a). The fact that the current spike returns to the baseline also suggests that it is likely due to a self-limiting process, which supports this hypothesis.

Frédéric Kanoufi enquired: You showed a very promising instrumental improvement for increasing the frequency range of single entity electrochemistry studies. It seems in the bandwidth spectrum curve presented in Fig. 2 of the paper (DOI: 10.1039/d1fd00055a) that at frequencies in the 10^5 Hz range (10^2 kHz) the gain is (slightly but still by >10%?) higher than at lower frequencies. Does it mean that in this frequency range, because of higher gain, the signal is overestimated compared to the other frequency ranges? If yes, and even if it is likely a minor effect, could you tell how much error can be made in terms of charge or current for quantitative single entity NP analysis?

Yi-Lun Ying replied: The slight increase in gain at high frequency in Fig. 3c of the paper (DOI: 10.1039/d1fd00055a) is caused by a little overshoot in the compensation circuit. We can solve this problem by adjusting the parameters of the compensation circuit. But it hardly causes an impact on the detection of the single entity in our experiments since the low-pass filter is set to 5 kHz.

Xu Liu communicated: The low-noise four-channel electrochemical instrumentation as described in your paper (DOI: 10.1039/d1fd00055a) looks promising. You talked about the challenges in achieving large-scale electrochemical measurement due to the current circuit designs with commercialized amplifier chips. Can you include more discussions on other factors (in addition to the circuit design) that could compromise the large-scale electrochemical measurement? Also, would you be able to comment on the potential of large-scale electrochemical measurement for other methods in the literature?

Yi-Lun Ying communicated in reply: To achieve large-scale electrochemical measurement, a parallelized working electrode is a must. The direct integration of the nanoelectrode array with the amplifier array is necessary to further reduce the capacitance from the electrode surface and connection wires. The high throughput measurement brings a massive dataset. Therefore, an automatic and intelligent data analysis algorithm is required to real-time and accurately read single-entity characteristics from the massive current data. In addition, during the process of electron transfer measurement at each working electrode, the reference/ground electrode combines with the unique one. Therefore, the reaction at the reference/ground electrode would cause the cross-talk. This factor should be considered for the further improvement.

Xiangkun Elvis Cao communicated: A vital component of the multi-channel amplifier arrays, as shown in your study, is the ability to reduce noise. The paper (DOI: 10.1039/d1fd00055a) studied some approaches for noise reduction (*e.g.*, by reducing the diameter of the Au electrode in the array). Can you also comment on some other potential methods to suppress the noise?

Yi-Lun Ying communicated in response: As discussed in our paper (DOI: 10.1039/d1fd00055a), the noise contribution of the ultra-low current measurement is mainly from the electrode interface, electrode connecting wire, amplifier input and amplifier parasitic capacitance of the feedback resistor. Therefore, the noise reduction methods should suppress the noise contributions from the above aspects. First, the small size of the electrode array (*e.g.* nanoelectrode array) certainly could reduce the capacitance from the electrode surface leading to a low current noise level. Moreover, the direct integration of the amplifier array with the micro/nanoelectrode would significantly minimize the connection distance, thereby, reducing the noise. Utilizing the multi-channel amplifiers designed using CMOS (Complementary Metal Oxide Semiconductor) technology could further suppress the noise. In addition, the current fluctuations from the charge carriers binding/unbinding with the sensing interface give the noise at the low frequency region. Therefore, the inhibition of the unwanted charge fluctuations (*e.g.* wetting and de-wetting, protonation and deprotonation) would benefit the noise reduction. The possible methods include the hydrophilic treatment (*e.g.* O_2 plasma) of the nanoscale sensing interface and the surface modification of hydrophilic compounds for the solution assay.

Serge Lemay opened discussion of the paper by Luke Gundry: This is a rather open-ended question, but thinking of a single-variable, quasi-continuous function $I(t)$ in terms of a 2D array of independent pixels takes some getting used to! To what extent does the coarse-graining process in two dimensions impact the information content of the CVs? Intuitively one expects that a fine enough grid will capture all of the information (our digitally recorded CVs are also discretized, after all), but I find it difficult to shake the feeling that *something* must be lost during this transformation.

Luke Gundry replied: Hey thanks for the question, its quite a technical one so my apologies in advance. What helped me conceptualize the 2D array was saving the 100 × 100 image to an Excel sheet where each pixel contains the value of the array. To imagine this simply think that the deep neural network (DNN) is “looking” at the voltammogram through squinted eyes. As for using a larger voltammogram 2D array, there are two main issues. The first issue is that the increase of array size drastically increases computational complexity (due to the n^2 relationship for a $n \times n$ image), the second issue is that the DNN is “looking” at the image with a kernel that scans across the 2D array for each layer. Where this kernel would have to be increased such that the DNN looks at the same amount of information in the comparative layers. Each layer also has a max-pooling after it which decreases the size of the array for the next layer, so the expansion would only matter for the first few layers anyhow.

Ultimately, a range of things were tested in this study which were not published, such as a 224 × 224 2D array with a more complex DNN layer design using the single and multiple channels. This approach is the more common in most image based DNNs, the biggest issue we had to contend with for these expanded DNNs was over fitting in the training process leading to poor generalization of the DNN, which could be corrected using more dropout layers but it was very challenging to get right for all reaction mechanisms we were originally trying. It was found that the simpler DNN with the inclusion of the 3 channels leads to the largest increase in improvement and didn’t have the issue with over fitting.

Overall, as much as there is definitely information lost in the coarse-graining process there is still enough information for the DNN to make an estimated prediction on what the reaction mechanism is for these reaction mechanisms.

Steven Linfield enquired: Can the DNN incorporate the mischaracterisations observed during testing with simulated data? For example, you demonstrated that the MCDNN could correctly assess 98.6% of the EC_1 simulations. Can the voltammogram features (or even simulation parameters) of the 1.4% of mischaracterised simulations be analysed and fed back into the DNN to improve the training or even shorten the training time? If this is possible, is there a limit to repeating this with less certain assessments (*i.e.* 84%) until probability of correct characterisation is high enough for real data?

Luke Gundry responded: This question was answered in the talk but I will add the plots used in the training of the DNN. In training, the parameters in the DNN are modified to maximise accuracy (a ratio of correct predicted mechanism for the training data) in the prediction and minimise the loss (a function to measure the error in the predictions) of the DNN predictions on the training data. Once this process has been completed for all the training data the DNN will then test itself on the remaining validation data. The full process is referred to as an epoch, where a DNN can complete multiple epochs in a row incorporating the errors in the prediction; the DNN should do as many epochs as is required for the accuracy to converge to a maximum and loss to converge to a minimum.

Fig. 2 here displays the model loss and accuracy plots for the multichannel DNN to show how the two values converge to a stable solution after 5 epochs.

Steven Linfield asked: You highlighted the difficulties of applying DNNs to the analysis of real voltammetry. Are there any downsides to including more simulation parameters (you mentioned for example convection, radial diffusion, and adsorption) that might model real experiments more closely and train the DNN for better characterisation of real voltammograms?

Luke Gundry answered: The main downside with including more parameters and increasing the complexity in the model is time. As in the computational time required to run the simulation and the manual time required to set up the mathematics in the simulation. For the simulation package used, convection can be modelled so including this would only lead to an increase of time to generate the simulation. However, since ten thousand simulations are required for each reaction mechanism a small increase in simulation time adds up quickly. Inclusion of say radial diffusion would lead to a massive increase in the mathematical modelling complexity and therefore time. This is due to the change from linear to radial coordinates so simulation of this would probably be unfeasible unless access to large computer infrastructure was available.

Stefano Fornasaro enquired: Could you please comment about the validation of the DNN models. How do you select the 80% of the data for training? I was wondering if you consider methods to check the instability coming from this choice.

Luke Gundry responded: The 80% for training is randomly selected from the 5004 for each reaction mechanism. Generally the training data needs to be enough samples to cover all possible cases which is why we have to simulate the large number of simulations for training. The other 20% of data is used for prediction validation and is the data not selected for training. Besides the optimisation process done in training which is discussed in another question, the instability of the prediction is checked to not be an effect by increasing the amount of training data used and generating a few DNN models using the same process to make sure it has converged to a constant rate of predictions and the

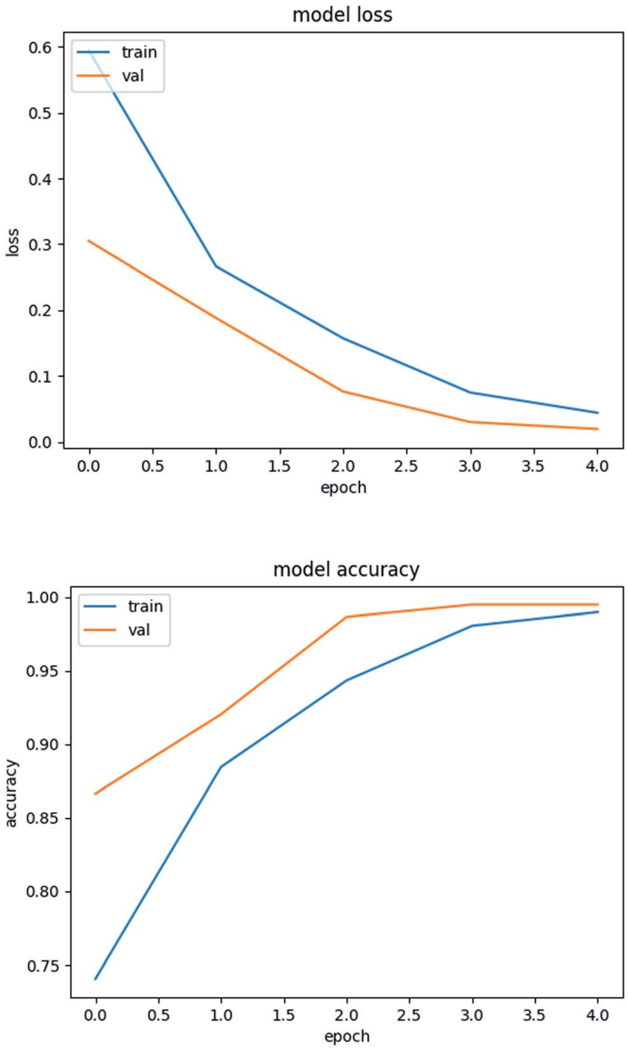

Fig. 2 Model loss and accuracy plots for the multichannel DNN.

predictions have converged to a stable answer independent of the randomly selected training data.

Tim Albrecht said: Following on from Fornasaro's question, in your work, you do not seem to use an "unseen" testset to test how well the network has learned. Did you optimise any hyperparameters during the validation phase? If not, it would seem that the validation set has been used as a testset. Is this correct?

Luke Gundry replied: For the DNNs published in this paper (DOI: 10.1039/d1fd00050k) we did not do hyperparameter tuning, the main reason for this was it was done under the study published by a coauthor of this paper.[1] Generally, I did try out a range of other hyperparameter values and none really led to any

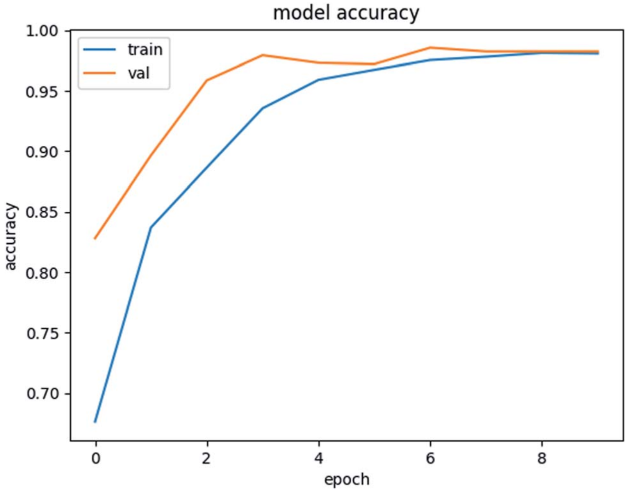

Fig. 3 Accuracy plot for a multicycle DNN.

improvements. For the training data themselves, we simulated more then enough data then split 80% of it for training and 20% for validation which was completed the end of each epoch, where 5–10 epochs were found to be required. For each epoch the DNN "looks" at all these simulations. Fig. 3 here displays an accuracy plot for a multicycle DNN and it should show that the accuracy steadily improves for each epoch, converging to a consistent value at around 5 epochs.

1 G. F. Kennedy, J. Zhang and A. M. Bond, *Anal. Chem.*, 2019, **91**, 12220, DOI: 10.1021/acs.analchem.9b01891.

Stefano Fornasaro added: Professor Albrecht is right. That procedure would result in a high risk of overfitting.

Frédéric Kanoufi asked: Thanks for this nice journey into artificial intelligence analysis of electrochemical data. You mentioned in your talk, a computer is not watching and analyzing an image (a cyclic voltammogram) as a trained scientist. As your article (DOI: 10.1039/d1fd00050k) suggests that some cases are difficult to classify upon the strategy proposed (using successive cycles), I wonder if a more classical analysis based on variations of CV response at different scan rates would be more helpful?

As a follow up question, increasing the scan rate may indeed increase the noise (and decrease the result of the classification), then could a strategy rely on trying to remove noise (by standard smoothing procedure, *etc.*) from the data before analyzing the data by the neural network?

Finally some mechanistic studies also rely on the evolution of peak potential, width or peak current with scan rate. Instead of analyzing an image could an analysis of the signal by identifying the peak descriptors be possible? Would it be faster to perform than an image analysis, would it be more accurate?

Luke Gundry answered: I believe I answered this question partly in the talk but I will add a few things. For the first part of the question the multiple scan rates or concentrations could be used, as long as it is constant over all simulated data and experimental input. In general these DNNs are very robust and flexible, the difficult part is simply making sure that the process is constant over all the data. As such, any form of data smoothing could be done to the image and I believe a moving average was used in the experimental data we tried. For the final part of the question it would be really hard to predict performance without attempting it as we simply do not know. Applying say an algorithm at simulation that identifies the peak, extracts them and then passes it to the DNN as a normalised 1D array could be possible, but it may bias the outcome due to the data extraction method, which is something you need to be wary of designing these algorithms.

Gabriel Meloni said: At the nanoscale EC mechanisms are usually poorly developed/hard to access. Piggy backing on Frédéric's point, maybe it would be interesting to investigate the scan rate dependency on $E_{1/2}$ for radial diffusion systems (similar to RDE studies for EC) to recover mechanistic information from nanoscale voltammograms. Would that be a possibility?

Luke Gundry responded: It may be possible but it is hard to say, there are two main issues with the suggestion.

One would be simulating all the training data which would be very time consuming due to the radial diffusion and the very high simulation accuracy required due to the very small scale affecting partial differential equation stability. This issue though time consuming can be worked through and with increased use of highly parallel cloud computing methods the time of one simulation would have less effect as you could run say 64 simulations at once.

The second issue is would the information be enough to make the labels in the system distinct enough? In a very recent paper from the Compton group[1] that uses a CE system, specific thermodynamic parameters can be identified. Though different it is a steady state system and they use a single voltammogram, which would suggest that your query could be applied. However, it would require a lot of independent simulations per training point increasing the number of total simulations by a multiple of the number of scan rates required. It would ultimately be possible though time consuming and not feasible at the current point. If there was a way to get more sensitive information from a single experiment, say by using AC voltammetry to get the systems more distinct over a single experiment.

1 H. Chen, E. Kätelhön, H. Le and R. G. Compton, *Anal. Chem.*, 2021, **93**, 13360, DOI: 10.1021/acs.analchem.1c03154.

Tim Albrecht enquired: Your work raises an interesting question about what the network actually learns. For example, increasing the potential window would only add capacitive charging and would not affect the relevant mechanistic information content in the data. Have you investigated this and could you comment on whether the network learns mechanistic information or rather simply what the image looks like? In the latter case, the result should be much more dependent on the exact scaling parameters, whereas in the former case this presumably would be less so.

Luke Gundry replied: Yes, we have tested this effect but its quite hard to pin down due to how the DNN looks at the image, due to the image using a moving kernel of a 3 × 3 array to identify features. If the potential range is changing the image will be compressed down to the size used and then "looked" at by the same 3 × 3 array. This array will then miss a lot of information as it will be different to what it's been trained for. The general question is does the DNN look at what it should be looking at as a chemist, to this extent the feature maps shown in Fig. 5 of the paper (DOI: 10.1039/d1fd00050k) shows it looks at the general image, with the main issue being that it looks at regions with more data points with an emphasis, which contains mechanistic information. This could be fixed quite easily by adding a weighted array that identifies the redox region and changes everything else to zero. Generally, the addition of three cycles adds more mechanistic information which is useful for the case explored in this paper.

Stefano Fornasaro queried: Is the code already available on Github?

Luke Gundry answered: Not as of the moment, DNN architecture can be seen in the supporting information of ref. 1. However, in future the code will be made available on my Github profile (https://github.com/lukegun) once our group have completed some things we are currently working on.

1 G. F. Kennedy, J. Zhang and A. M. Bond, *Anal. Chem.*, 2019, **91**, 12220, DOI: 10.1021/acs.analchem.9b01891.

Dongfei Chen asked: Is it possible to use your method to perform baseline subtractions for square wave voltammograms or differential pulse voltammograms?

Luke Gundry replied: Its difficult to say without knowing more about the system, a Deep Neural Network (DNN) could probably be trained for baseline subtraction. There was a really good paper by Dean *et al.*[1] that used square wave voltammetry and deep neural networks for chemical identification, which is of a similar vein.

I will say that due to the computational time required for DNN training it may be advantageous to try a more direct method first, such as using the first and second derivative of the current for region identification. Feel free to email me further if you want to discuss a more specific application.

1 S. N. Dean, L. C. Shriver-Lake, D. A. Stenger, J. S. Erickson, J. P. Golden and S. A. Trammell, *Sensors*, 2019, **19**, 2392, DOI: 10.3390/s19102392.

Grant Jeffcoat commented: As I understand by reading the paper; the neural networks struggled with voltammetric results that even humans would struggle with. And as you pointed out, the networks are trained on idealized data from models.

Based on this, I was wondering what you envision for immediate applications of this software. Additionally, what kinds of applications do you foresee with future development. You did outline some possible development for the networks. How might these expand or open applications?

Luke Gundry responded: Our group discussed the future application more in a recent *ChemComm* paper.[1] However, the current application of the code presented in the paper would be minimal due to the difficulty in running the code for a general chemist and the prediction made could simply be identified by doing the same experiment at a couple of different analyte concentrations. The usefulness of these study is that inclusion of different channels for a DNN helps its prediction capabilities on tougher systems. Future aspects of this kind of technology would be useful, as testing experimental data is very quick and can be designed to be very general for a wide range of reaction mechanisms. As such, an initial reaction mechanism could be predicted qualitatively in seconds before passing it to more advanced/longer running quantitative algorithms as shown in Fig. 4 here, which is taken from ref. 1.

1 L. Gundry, S.-X. Guo, G. Kennedy, J. Keith, M. Robinson, D. Gavaghan, A. M. Bond and J. Zhang, *Chem. Commun.*, 2021, **57**, 1855, DOI: 10.1039/d0cc07549c.

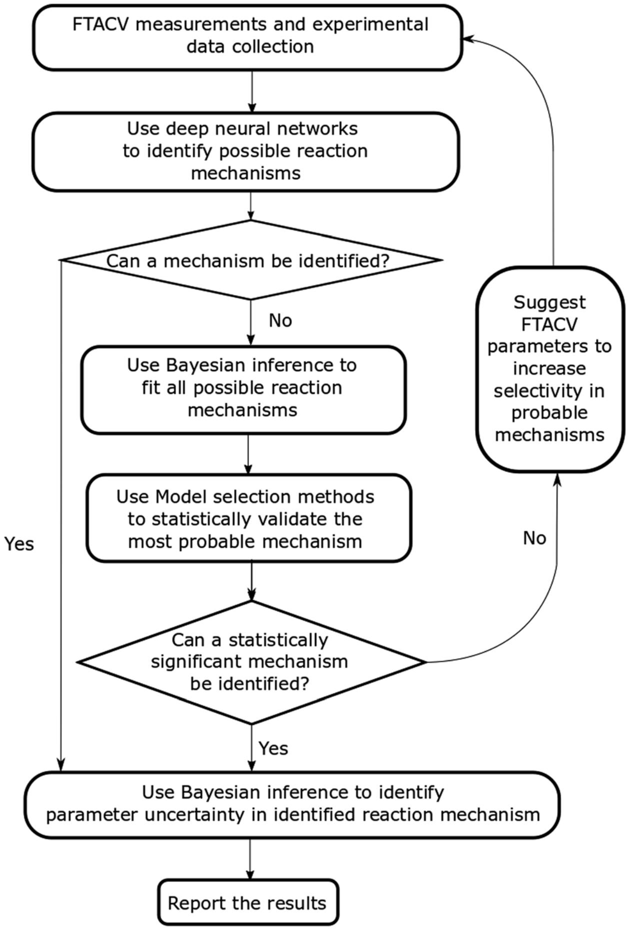

Fig. 4 Proposed artificial intelligence flowchart for analysis of FTAC voltammetric experimental data. Reproduced from Ref. 1 with permission from the Royal Society of Chemistry.

Martin A. Edwards communicated: Thanks for the interesting talk in an exciting area. In your talk, you alluded to one of the next steps for this work being expanding this to a wider range of mechanisms. I am unfamiliar with the networks you use and am curious about how such a system is anticipated to scale with the number of possible outcomes. Would you anticipate a network of similar complexity/depth to have sufficient power to resolve more mechanisms? Or would you have to keep adding layers to achieve comparably good resolution? I presume that this makes it slower/harder to do computationally, but have no idea of the anticipated scaling in the size/complexity of the network and increase in training time with number of outcomes. Do you have an idea how quickly each of these might grow? Does the current approach seem plausible for the number of mechanisms you plan to look at, or do you anticipate having to add something extra to make that step?

Luke Gundry communicated in reply: Originally the DNNs were tested on 7 reaction mechanisms and performed reasonably decently, the largest issue is the generalization when applied to out of sample data/experimental data and there were issues with differentiating between systems which were visually identifiable. In a similar vein to a lot of Savéant's work, there are a lot of voltammograms that can be fit by two different reaction mechanisms depending on the parameters used. Understandably, in this case voltammograms can't differentiate between the two reactions, with the most apparent case of this being E and EE mechanisms where the two electron transfers overlap. To work around this generally smart parameter ranges need to be selected to differentiate identical systems, an automated way to classify these labels would be good. Overall, to get a decent value and good differentiation of reaction mechanisms (as well as identify quantitative values) a range of algorithms working together will be required to correctly differentiate all possible cases.

Conflicts of interest

There are no conflicts to declare.