



Smoothing Methods for Numerical Differentiation to Identify Electrochemical Reactions from Open-Circuit-Potential Data

Mark W. Verbrugge,¹ Charles W. Wampler,² and Daniel R. Baker¹

Chemical and Materials Systems Lab, General Motors Research and Development, Warren, Michigan 48092, USA

We develop and compare methods to determine when electrochemical reactions take place within intercalation electrodes. We employ open-circuit data from the Chevrolet Bolt EV negative and positive electrodes, lithiated graphite and NMC ($\text{Ni}_{0.6}\text{Mn}_{0.2}\text{Co}_{0.2}\text{O}_2$), respectively. Electrochemical reactions correspond to peaks in plots of dx/dU vs. U , where U is the electrode potential and x is the fractional state of charge ($0 \leq x \leq 1$). We employ three nonparametric methods in this work to smooth and subsequently differentiate the data: a windowed cubic polynomial of fixed and variable window length and a cubic smoothing spline. The standard deviation in the measurements can be used to set appropriately the window length and the analogous smoothing parameter for the spline. The cubic polynomial with a fixed window length and the smoothing spline each have a single smoothing parameter; both methods are shown to work well and yield comparable results. The best results were achieved with the cubic polynomial incorporating a varying, adaptive window length, and the increased fidelity enables a completely automated procedure to identify reactions. The analysis clarifies the multi-reaction nature of lithiated graphite, which must be addressed in cell modeling to simulate fast-charge behavior. © The Author(s) 2018. Published by ECS. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 License (CC BY, <http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse of the work in any medium, provided the original work is properly cited. [DOI: 10.1149/2.0951816jes]



Manuscript submitted October 10, 2018; revised manuscript received December 13, 2018. Published December 27, 2018.

The importance of differential voltage spectroscopy (DVS) is well explained in three recent publications. Pastor-Fernández et al.¹ overviewed the utility of incremental capacity-differential voltage (termed DVS in this work³) to characterize degradation phenomena in battery systems. The authors characterize the state-of-health of the batteries by examining the shift in peaks obtained from differentiating the open-circuit voltage (OCV) curves $U(x)$ at low, constant-current operation ($C/25$ in their investigation, where the 1C rate corresponds to the discharge current that will discharge the cell in 1 hour). Related to work described in Ref. 1, the utility of DVS for informing life models of lithium ion batteries, including peak shifting associated with coupled chemical degradation and fatigue mechanics, is examined in detail for a graphite/iron-phosphate cell in Ref. 2. When such detailed model-experiment comparisons are made, it is helpful to smooth the data consistent with the amount of noise, and thereby enable the robust construction of plots requiring the differential dx/dU (or the inverse, dU/dx). The smoothing process facilitates the fitting of model parameters to the differentiated data. Li et al.³ review the DVS literature and apply differentials of $U(x)$ curves for deducing state-of-health and state-of-charge of lithium ion cells; they employ a Gaussian filter to smooth the $U(x)$ curves. At this time, it remains an open question as to the merits of a Gaussian filter relative to those of the methods described in this work. Last, Zheng et al.⁴ review the DVS literature and, like Refs. 1 and 3, emphasize the utility of smoothed $U(x)$ curves for deducing the state of lithium ion cells and algorithms derived by modeling the differential data for battery management systems.

While smoothed $U(x)$ curves are useful for the reduced-order modeling for estimating state-of-health and state-of-charge, they are also important for the modeling of porous electrodes with transport relations based on irreversible thermodynamics.^{5,6} We can write the flux of lithium-filled sites N within an insertion or alloy electrode as^{7,8}

$$N = c_T x(1-x)D_0 \frac{F}{RT} \frac{\partial U}{\partial x} \nabla x + x(N + N_H) \quad [1]$$

where N_H is the flux of vacant sites, c_T is the total concentration of sites available for lithium, and D_0 is the diffusion coefficient of lithium in the limit as x tends to zero (a condition under which Fickian diffusion prevails). Faraday's constant, the universal gas constant, and

absolute temperature are given by F , R , and T , respectively. From Eq. 1, we can see that an accurate model is needed for dU/dx if the flux expression is to be accurate.

In addition to describing lithium diffusion within insertion or alloy electrodes, it is important to know how many electrochemical reactions are needed to represent electrode behavior, a fact that is often overlooked in the modeling of intercalation electrodes, as it is common to assume that a single electrochemical reaction, described by a Butler-Volmer equation, prevails: this approach conflicts with the observation of multiple peaks observed in DVS. A model incorporating multiple lithium species and associated reactions has recently been developed¹⁰ to rectify this situation. The different species of lithium correspond to different galleries in the host materials and were shown to be consistent with X-ray analyses. The single Butler-Volmer equation must then be replaced with an electrochemical reaction expression for each lithium species determined. Here again, an accurate determination of dx/dU from DVS data plays a critical role in the process. For linear-sweep voltammetry,⁹ current is proportional to dx/dU at low currents; hence, we can associate peaks in DVS plots (dx/dU plotted against U) with current peaks in voltammograms (plots of current vs. potential, with the potential being scanned linearly with time). One must ascertain how many significant peaks there are in the DVS plots, corresponding to electrochemical reactions, to model accurately voltammetry experiments⁹ using information obtained only from DVS experiments. More discussion on the similarity of DVS at very low currents to cyclic voltammetry at low potential scan rates can be found in Ref. 10. In this context, Figure 1 provides a schematic overview of one of the objectives of this work. To identify electrochemical reactions, we determine when $d^2U/dx^2 = 0$, as will be described, requiring a second numerical differentiation of the data, and placing a premium on smoothing the data but retaining sufficient fidelity in the data features.

This work is concerned with the process of determining a differentiable function $U(x)$ from DVS data. To do this, one must filter

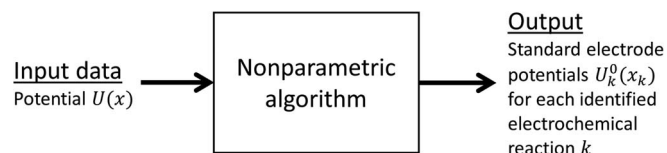


Figure 1. Schematic overview of the objective of this work. Given the input data $U(x)$, provide a feature-recognition (nonparametric) algorithm to identify electrochemical reactions, characterized by standard potentials $U_k^0(x_k)$ for each reaction k .

¹Electrochemical Society Fellow.

²Electrochemical Society Member.

³E-mail: mark.w.verbrugge@gm.com

⁴We examine dx/dU vs. U in this work, but the same methods apply for analyses of dx/dU vs. x , dU/dx vs. U , and dU/dx vs. x , as the key challenge is accurate numerical differentiation of $U(x)$ curves.

out noise in the data, because unfiltered noise impacts not only the function $U(x)$, but its derivatives even more so.

Methods for analyzing experimental data can be classified as *parametric* or *nonparametric* according to whether or not they are based on a model that depends on physicochemical parameters of the system being studied. When a well-established mathematical model is available that captures the significant physical phenomena occurring in the experiment, it is appropriate to regress unknown parameters in the model. Good agreement between the regressed model and the experimental data lends credence to the belief that the model is appropriate, and studies of the sensitivity of the experiment to the parameters can help quantify the confidence in the parameters found by regression. However, if one applies an inappropriate model, especially if that model is “over-parameterized,” a good agreement in the regression might still be possible, leading to a false belief in the model that may not stand up when faced with new evidence.

Nonparametric methods stand in contrast to the parametric approach. A nonparametric method attempts to extract functional relationships from the data using minimal prior information, typically limited to basic assumptions on the smoothness of the functions and the accuracy of the experimental apparatus. Supposing that two experimental variables have a functional relation, say $y = \mu(x)$, Eubank¹¹ offers the following perspective:

“... the use of parametric methods can be quite dangerous in situations where there is little known about the regression function. In such cases, the bulk of the information about μ lies in the data rather than the person conducting the study or experiment. Accordingly, it seems more reasonable to use inferential techniques which rely heavily on the data. It is for this reason that nonparametric regression techniques are ideally suited to problems of inference when the available knowledge about μ is limited in nature.”

In the work presented here, we analyze experiments on battery electrodes and use nonparametric methods to extract a smooth open-circuit voltage curve and its derivative. Features of the smoothed function identify where electrochemical reactions occur as the electrode is charged or discharged. In previous work,^{10,12} a parametric model was used to fit the open-circuit voltage curve to data, requiring prior knowledge of how many reactions are occurring. The methods discussed here require no such prior knowledge. As discussed above, the nonparametric methods provide a basis for using the parametric model and assessing its accuracy.

Intercalation and alloy electrodes that contain many phases are difficult to model, and many parameters appear in, for example, porous electrode models.^{5,6,13} Recently, the difficulty in isolating meaningful values led us to perform a sensitivity analysis to examine the impact of the 20 physicochemical parameters used in the modeling of a lithium.¹² More information that allows for clarification of individual electrode reactions would be of high value in terms of formulating models for porous electrodes comprising intercalation and alloy active materials. It has been previously shown (see Ref. 10 and references therein) that reactions in electrochemical systems can be identified by differentiation of the potential U of an electrode (relative to that of a reference electrode) with respect to composition (e.g., fractional state of charge x , with $0 \leq x \leq 1$); i.e., plots of dx/dU vs. U , as embodied in the terminology “differential voltage spectroscopy” (DVS). For the extraction of dx/dU from the common experiments that involve a low but constant current, numerical differentiation of the potential trace is required. Thus, we seek to understand features in plots of dx/dU vs. U , and that is the subject of this work.

Differentiation of data is a challenging endeavor that has a long history. The most common approach can be broken down into two parts: first, a function is fit to the data, and, second, the fit function is differentiated. Polynomial regression estimators are routinely employed as fitting functions. Eubank (section 3.5 of Ref. 11) reviews the history of polynomial regression and points out the motivation for the method based on theoretical underpinnings: Taylor’s Theorem

and the Weierstrass Approximation Theorem can be used to bound the error of a finite order polynomial regression estimator. This does not mean the error will be small, however, particularly for a lower order polynomial estimator. Regression with a lower order polynomial (e.g., a cubic polynomial) can be expected to work well for low-frequency components of data, and, if the data is primarily composed of low-frequency information, a smooth and useful fitting function is rendered. Attempts to modify lower order polynomial estimators to address the remainder terms (e.g., high-frequency components) have led to the development of smoothing splines^{11,14,19} and least-squares spline estimators.^{11,19} Conte and de Boor quantify the error in the differentiation of a cubic interpolation spline, and they specifically underscore the utility of such splines for differentiation of data (section 6.7 of Ref. 15). Richardson et al.¹⁶ show the utility of a cubic interpolation spline for machine learning purposes with comparison to neural networks and fuzzy systems. Feng¹⁷ discusses the utility of a cubic smoothing spline (for equally spaced data points) in the context of data differentiation and shows that a smoothing spline is “a lowpass filter with the maximum flatness property,” a statement that is further explored in that work. Russell and Norvig¹⁸ overview the use of nonparametric regression in the context of agent (or machine) learning, which requires the repeated use of feature recognition.

This short perspective on the differentiation of data, and the need to determine dx/dU for electrochemical systems, such as battery electrodes, motivate the application of nonparametric regression to fit $U(x)$ based on (1) a windowed cubic polynomial estimator (including the use of a variable window length of data) and (2) a cubic smoothing spline. The resulting regressions are then differentiated to estimate dx/dU . We employ the open-circuit potential data of Ref. 10 for lithiated graphite and the lithiated metal oxide $\text{Ni}_{0.6}\text{Mn}_{0.2}\text{Co}_{0.2}\text{O}_2$; both are relevant materials that are employed in the Chevrolet Bolt electric vehicle. Accurate modeling of electrode systems, such as graphite and $\text{Ni}_{0.6}\text{Mn}_{0.2}\text{Co}_{0.2}\text{O}_2$ that are used today, is key to designing batteries and control systems that can accommodate faster charge rates and avoid unwanted side reactions, such as lithium plating on overcharge.²⁰⁻²² We also discuss and formulate methods to determine how many points should be included in the cubic window regression (cf. Eqs. A7 and A10 and for the determination of L for a constant window width and L_i for an adaptive window width, respectively), and the smoothing parameter p for the smoothing spline (cf. Eq. B23).

Numerical Methods

We provide the numerical details for the implementation of a windowed cubic polynomial estimator (including the use of a variable window length of data), and in Appendix A, we show how to apply a cubic smoothing spline in Appendix B. Common to all methods examined in the work is the use of cubic polynomials combined with a penalty on curvature in the smoothing functions in a manner that scales with the noise in the measurements. For the windowed cubic polynomial estimator with a fixed length of data points L , Equation A7 provides a relationship to determine the window length consistent with the noise in the data σ . Equation A10 provides the corresponding relationship for the variable window length estimator. For the smoothing spline, Equation B23 provides a relationship to determine the smoothing parameter p consistent with the noise in the data σ . We shall describe in the next section of this paper how we estimate σ (cf. Figure 3 and the related discussion).

We are not aware of a publication treating a cubic polynomial estimator with a variable window length of data, including the head, middle, and tail regions of the data (cf. Algorithm 2 in Appendix A). For the smoothing spline, the incorporation of proper scaling has not, to our knowledge, been reported previously (cf. Eqs. B9 and B10).

Results and Discussion

Cubic polynomial regression over a constant-width moving window for lithiated graphite.—The application of the cubic fit with a constant-width window (cf. Algorithm 1 of Appendix A) is shown in

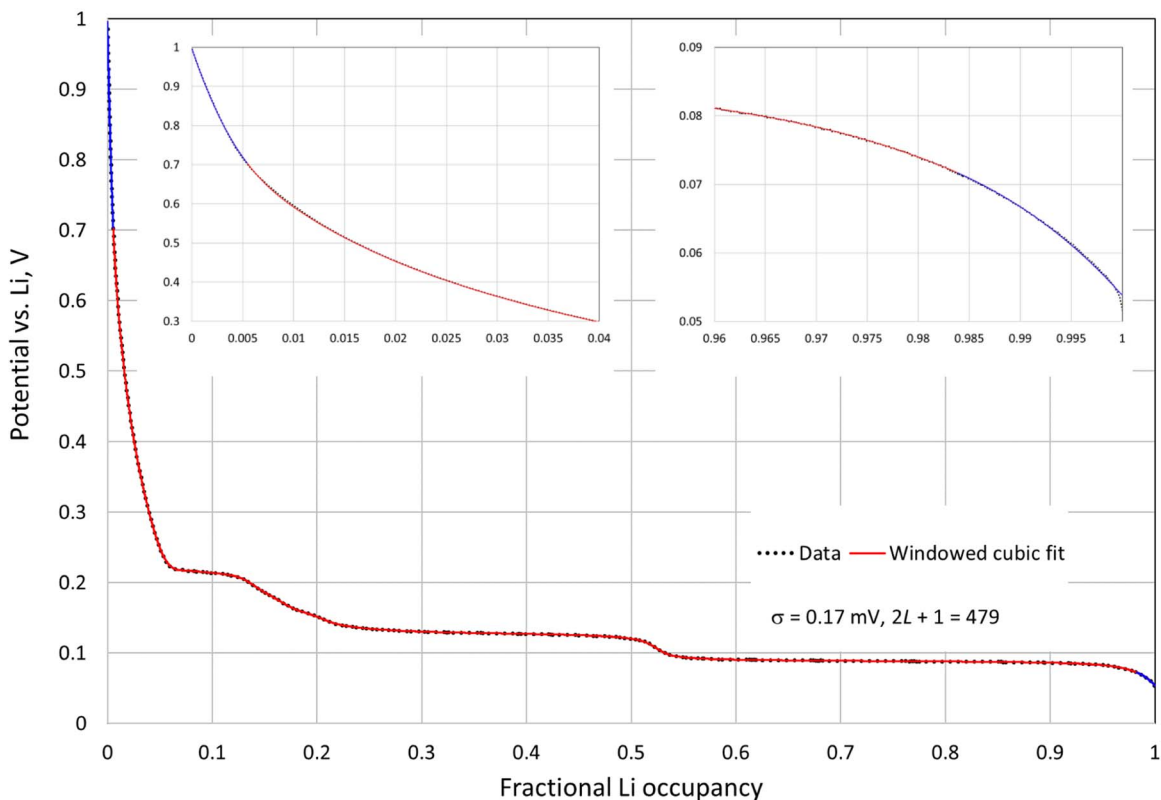


Figure 2. Cubic polynomial fit to the graphite data based on $U(x_i)$ approximated by Algorithm 1 of Appendix A, where $\sigma = 0.17$ mV is consistent with $2L + 1 = 479$ points for the constant-width moving frame over which the cubic polynomial is regressed.

Figure 2. The inset plots of Figure 2 show blue curves for the first and last of $L = 239$ data points. We exclude data below 80 mV and above 250 mV for calculating the sum of squared residuals in Eq. A6, as the slope in the actual $U(x)$ data is significant below 80 mV and above 250 mV, so we don't penalize curvature for those potential ranges. For the graphite data, there were 14,132 total data points, with 12,692 within the range $80 \text{ mV} \leq U \leq 250 \text{ mV}$ (see the shaded potential regions that are excluded in the inset plot of Figure 3); hence, $N = 12,692$ in

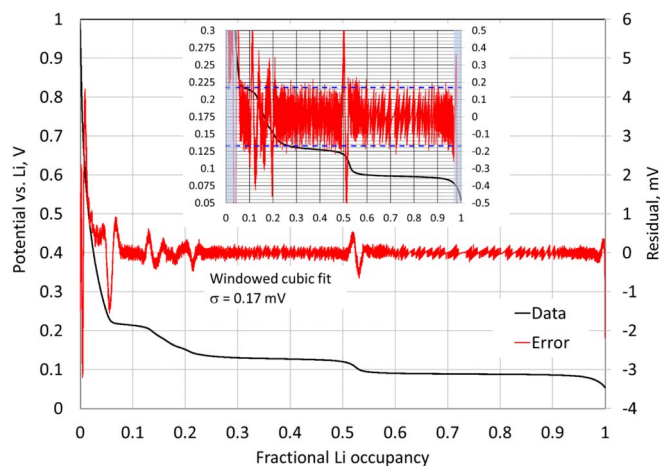


Figure 3. Residual in the cubic polynomial fit shown in Figure 4. Within the inset figure, the dashed lines correspond to ± 0.17 mV, consistent with $\sigma = 0.17$ mV, and the shaded portions indicate voltage ranges (below 80 mV and above 250 mV) excluded from use in calculating the squared residuals to determine L (cf. Eq. A6).

the use of Eqs. A6 and A7 to determine the window length using the standard deviation $\sigma = 0.17$ mV (to be discussed below).

Shown in Figure 4 are expanded views of the potential traces along two of the potential plateaus of Figure 2, which occur near 89 and 215 mV. From this plot (and many others not shown in this publication), we find that the measurement error in our equipment is about 0.15 mV for the graphite experiments. Hence, we expect σ to be near this value. We differentiate the cubic regression to approximate dx/dU , $dx/dU \approx 1/(a_1 + 2a_2x + 3a_3x^2)$, and by plotting $U(x)$ versus dx/dU , we obtain the result shown in Figure 5. We find that window lengths less than $2L + 1 = 479$ data points did not provide

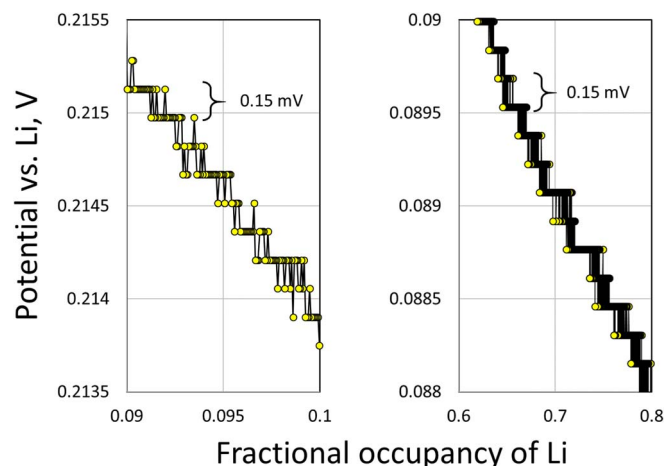


Figure 4. Noise in the experimental data on the approximately 89 and 215 mV $U(x)$ plateaus for the lithiated graphite.

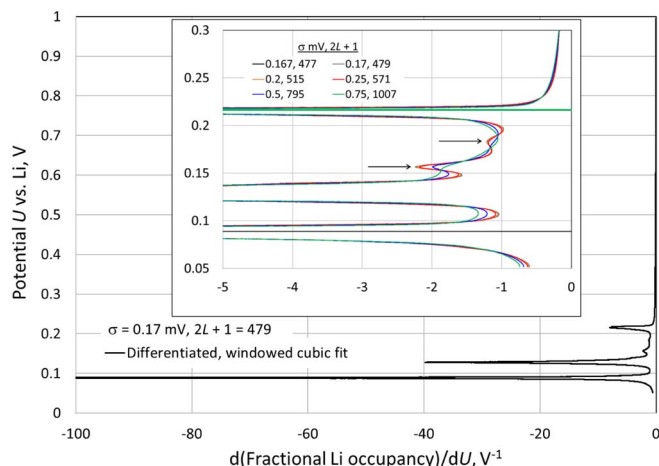


Figure 5. Differential voltage spectroscopy plot constructed with the cubic regression depicted in Figure 2. The inset plot shows influence of changing σ , and thus L (cf. Eq. A7).

enough smoothing, with the result that dx/dU at 89 mV became positive, which is not physically possible (batteries do not increase in potential energy upon discharge, and dx/dU must always be negative for lithiated graphite). The inset plot of Figure 5 for $2L + 1 = 477$ shows that dx/dU transitions abruptly to 0 at 89 mV. (The maximum in the abscissa of Figure 5 is zero, and the positive values for dx/dU are not depicted as plotted.) When $2L + 1$ is set to 479 (i.e., L is set to 239), we find from Eq. A6 that $\sigma = 0.17$ mV. From this observation, and that of σ being near 0.15 mV per Figure 4, we conclude that a value of $\sigma = 0.17$ mV is appropriate. The influence of changing σ , and thus L , is depicted in the inset plot of Figure 5. Larger values of σ and L give rise to more smoothing and loss of peak and valley amplitudes. For σ greater than about 0.25 mV, one can no longer see the small peak associated with the uppermost arrow. For σ greater than about 0.75 mV, the smoothing is excessive, and aberrations evolve, such as the positive value for dx/dU at about 89 mV, which, as noted previously, is not physically reasonable.

Smoothing spline regression for lithiated graphite.—For treating lithiated graphite with a smoothing spline, we determine the appropriate level of smoothing using the same settings as was done for the windowed cubic polynomial regression, that is, (i) we use $\sigma = 0.17$ mV for the noise in the voltage measurements and (ii) we penalize curvature only for data in the potential range $80 \text{ mV} \leq U \leq 250 \text{ mV}$, leading to $N = 12,692$. For that range, the scaling factor s from Eq. B10 evaluates to 17.92 mV, which with $x_N - x_1 = 1$ and $\sigma = 0.17$ mV, leads to $N(x_N - x_1)^3 \frac{\sigma^2}{s^2} = 1.137$ in Eq. B12. This happens to be close to 1, so we simplify Eq. B12 to

$$\lambda = \frac{1-p}{p} \text{ or } p = \frac{1}{1+\lambda} \quad [2]$$

as $N(x_N - x_1)^3 \frac{\sigma^2}{s^2}$ is sufficiently close to 1 for this work. When we specify the value of p in the figures to be discussed, Eq. 2 is employed for λ , and the smoothing spline coefficients are calculated by means of Eqs. B18–B22.

The general behavior of the smoothing spline is shown in Figure 6. Different values of the smoothing parameter p , corresponding to different values of σ (cf. Eq. B23), are provided in the legend of Figure 6. Because $p = 1^-$ (p less than but tending to 1) corresponds to no smoothing, that extreme gives an interpolation spline that passes through each data point, i.e., $y_i = a_i = U(x_i)$ vs. x_i . The opposite extreme of $p = 0^+$ gives the least-squares best-fit line $l(x_i)$, which enters Eq. B10 for the calculation of $s = 17.92$ mV, the value mentioned above. Curves for values of p between 0 and 1 smoothly transition between the two limiting cases.

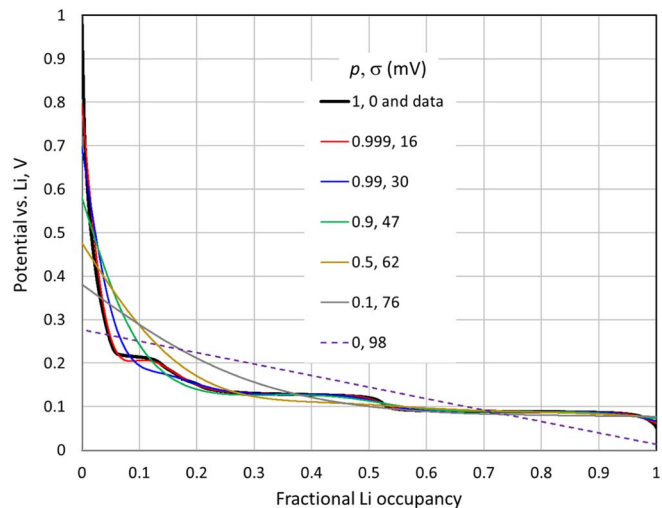


Figure 6. Impact of the parameter p on the smoothing spline. For $p = 1^-$ (i.e., p tending to but less than 1, corresponding to λ tending to 0), a cubic interpolation spline results, which passes through each data point. For $p = 0^+$ (λ growing to infinity), a least-squares fit of a line through the data results. As shown in the legend, a value of σ can be associated with each p (cf. Eq. B23). The implication of the p – σ relationship is that lower standard deviations in the measurements are to be associated with higher p values, with $p = 1$ corresponding to no noise in the measurement.

The plots in Figure 7 for the smoothing spline analysis of lithiated graphite are analogous to those of Figure 2 for the windowed cubic polynomial. For the smoothing spline, there are no end regions, in contrast to the treatment of the first and last L data points for the windowed cubic, but the errors are highest near the end points, as is evident in Figure 7 and Figure 8. The inset plot of Figure 8 shows an expanded view for both ordinates. Evident are the dashed lines for $\sigma = \pm 0.17$ mV and the shaded regions for potential outside the range of $80 \text{ mV} \leq U \leq 250 \text{ mV}$.

A differential voltage spectroscopy plot using the fit smoothing spline coefficient $1/b_i$ (see line 2 of Eq. B2) is provided in Figure 9. For the smoothing spline applied to the graphite data, we find that for σ values below 0.1 mV, positive values of dx/dU result near $U = 89$ mV; the results for $\sigma = 0.1$ mV and $\sigma = 0.17$ mV (the latter plotted in Figure 9) are difficult to distinguish, as are the results for

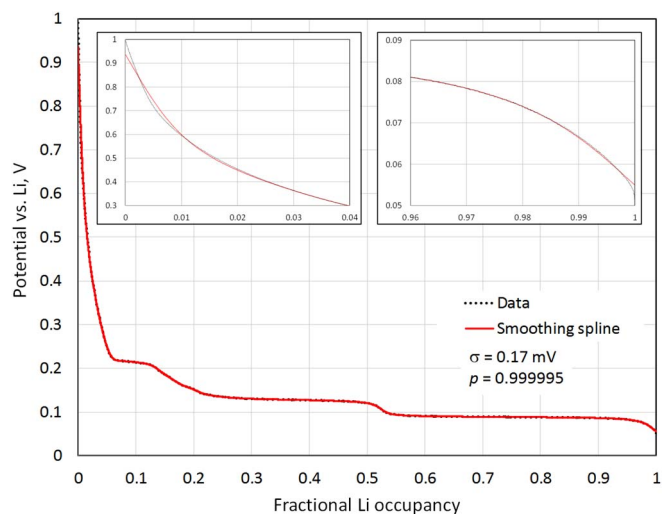


Figure 7. Smoothing spline fit to the graphite data based on $U(x_i)$ approximated by a_i (cf. first line of Eq. B2). Per Eq. B23, $\sigma = 0.17$ mV is consistent with $p = 0.999995$.

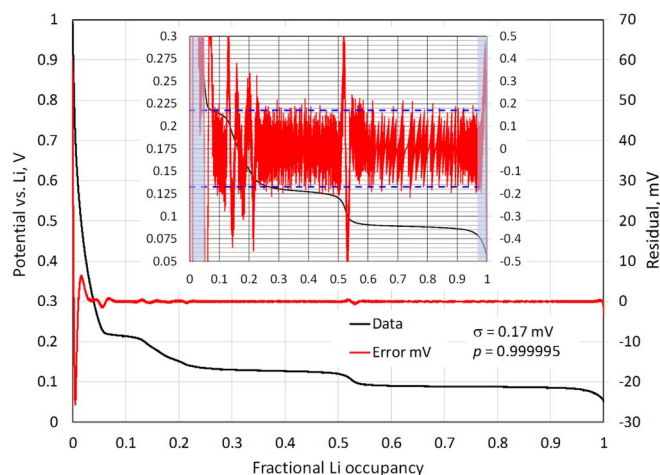


Figure 8. Residual in the smoothing spline fit shown in Figure 7. The residual corresponds to $U(x_i) - a_i$. Within the inset figure, the dashed lines correspond to ± 0.17 mV, consistent with $\sigma = 0.17$ mV, and the shaded portions indicate voltage ranges (below 80 mV and above 250 mV) excluded from use in calculating the squared residuals to determine p (cf. Eq. B23).

the windowed cubic and smoothing spline employing $\sigma = 0.17$ mV, both of which are displayed in the inset plot of Figure 9. As with the windowed cubic, for $\sigma = 0.75$ mV, dx/dU turns positive near $U = 215$ mV, indicating excessive smoothing (see the green curve in the inset plot of Figure 9).

In general, the results for the windowed cubic polynomial (Figure 2–Figure 5) are quite similar to those for the cubic smoothing spline (Figure 7–Figure 9). However, the relationships between the standard deviation in the data σ and the window length L and the smoothing parameter p are quite different when plotted as in Figure 10. The vertical dashed line corresponds to $\sigma = 0.17$ mV, $L = 239$, and $p = 0.999995$. For 0.15 mV $\leq \sigma \leq 0.8$ mV, L varies linearly, to good approximation, with σ , with L ranging from about 200 to 500 data points. Conversely, over the same σ -range, p changes over a very small range near unity (about 0.9985 to 1⁻), and the relationship is nonlinear over the range 0.15 mV $\leq \sigma \leq 0.8$ mV.

Smoothing spline regression for lithiated $\text{Ni}_{0.6}\text{Mn}_{0.2}\text{Co}_{0.2}\text{O}_2$.

Plots for the application of the cubic smoothing spline regression applied to lithiated $\text{Ni}_{0.6}\text{Mn}_{0.2}\text{Co}_{0.2}\text{O}_2$ can be seen in Figure 11 and

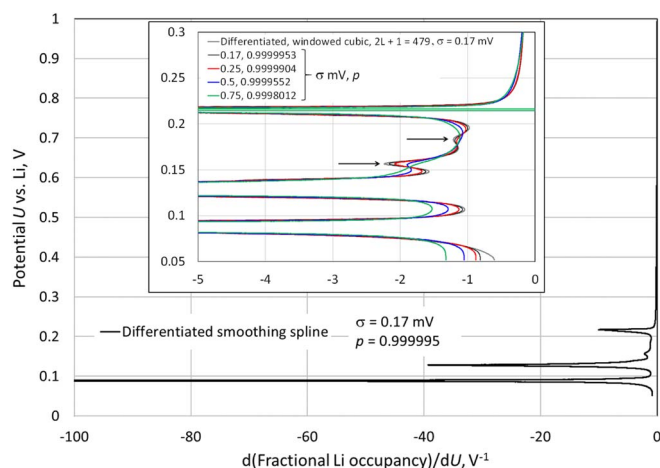


Figure 9. Differential voltage spectroscopy plot constructed with the smoothing spline depicted in Figure 7. Also shown is the differentiated, windowed cubic from Figure 5 ($\sigma = 0.17$ mV), which is closest to the cubic smoothing spline curve with $\sigma = 0.17$ mV.

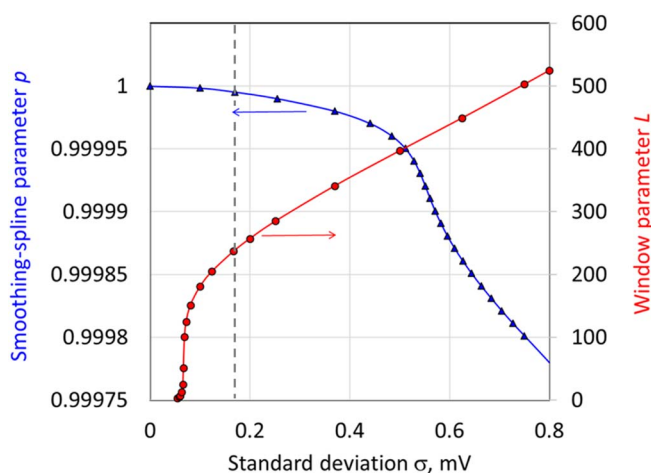


Figure 10. Smoothing spline parameter p and windowed cubic parameter L . For $\sigma = 0.17$ mV, $p = 0.999995$ and $L = 239$, corresponding to the dashed, vertical line and Figure 2 and Figure 7, respectively. Symbols correspond to calculations, and the solid lines are provided to guide viewing.

Figure 12. Equation 2 was employed for λ – p relationship. To avoid penalizing curvature where it is significant in the data, we do not penalize curvature for potentials below 3.6 V vs. Li. Figure 11 provides plots of U vs. x and U vs. dx/dU . The sensitivity of the U vs. dx/dU regression to the value of σ (and thus p) is shown in the left plot of Figure 12. The plot on the right shows the error in the regressed U vs. x curve relative to the data. The inset plot of Figure 12 shows an expanded view for both ordinates. Evident are the dashed lines for $\sigma = \pm 0.25$ mV and the shaded region for potentials below 3.6 V, which are not included in the calculation of the squared error term of Eq. B23.

A larger value for the standard deviation in the data σ was employed for the $\text{Ni}_{0.6}\text{Mn}_{0.2}\text{Co}_{0.2}\text{O}_2$ material (0.25 mV) vs. that for the graphite material (0.17 mV), consistent with our finding slightly larger variations in the NMC experiments, which can be seen by comparing Figure 4 with Figure 13. In both cases, the values of σ are small insofar as the 5V full-scale measurement range has an error of about 0.17/5000 or 0.0034% for the graphite measurements and 0.25/5000 or 0.005% for the NMC measurements.

Automated reaction recognition enabled by an adaptive window.—To fulfill the objective associated with the schematic in Figure 1, that is, to identify when reactions are taking place, we recognize that for a cubic regression:

$$\begin{aligned}
 p(x) &\approx U(x), & p(x) &= a_0 + a_1x + a_2x^2 + a_3x^3, \text{ and} \\
 p' &= \frac{dp}{dx} = a_1 + 2a_2x + 3a_3x^2 \\
 \frac{d^2x}{dp^2} &= \frac{d}{dp} \frac{dx}{dp} = \frac{dx}{dp} \frac{d}{dp} \frac{dx}{dp} = \frac{1}{p'} \frac{d}{dx} (p')^{-1} = -\frac{p''}{(p')^3} \text{ where} \\
 p'' &= 2a_2 + 6a_3x
 \end{aligned}
 \quad [3]$$

The quantity $1/p' \approx dx/dU$ is always negative. Hence, to deduce when a maximum or minimum occurs in $1/p' \approx dx/dU$, we can track $p'' \approx d^2U/dx^2$. A reaction is identified by a positive-slope crossing of d^2U/dx^2 through the line $d^2U/dx^2 = 0$, corresponding to a minimum (a reaction peak) in the dx/dU trace. A negative-slope crossing corresponds to a maximum (valley) in the dx/dU trace.

While the constant-width window for the cubic regression and the smoothing spline provide similar and useful results for representing $U(x)$ and dx/dU , the smoothing afforded by these two methods is insufficient for the lithiated graphite system for $p'' \approx d^2U/dx^2$ in the vicinity of the two largest peaks, as is seen in the left plot of Figure 14. (The results for the constant-width window for the cubic regression is effectively identical to those plotted in Figure 14.) The values for $p'' \approx d^2U/dx^2$ are sufficiently smooth to treat the data for the $\text{Ni}_{0.6}\text{Mn}_{0.2}\text{Co}_{0.2}\text{O}_2$ host material.

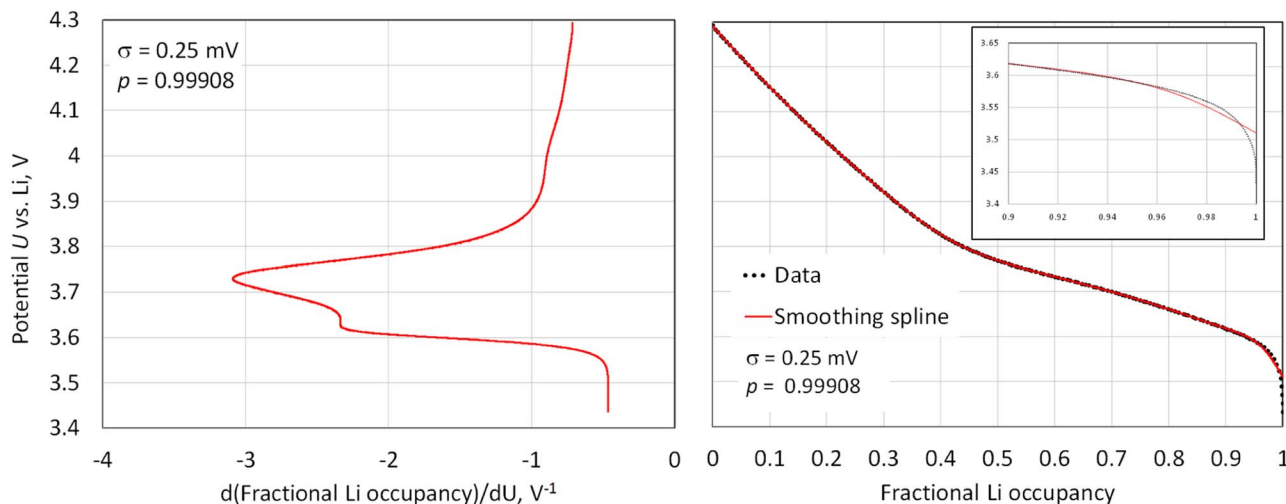


Figure 11. Smoothing spline fit to the $\text{Ni}_{0.6}\text{Mn}_{0.2}\text{Co}_{0.2}\text{O}_2$ data based on $U(x_i)$ approximated by a_i (cf. first line of Eq. B2), which is obtained ultimately through the use of Eq. B22. Per Eq. B23, $\sigma = 0.25$ mV is consistent with $\rho = 0.99908$ for the $\text{Ni}_{0.6}\text{Mn}_{0.2}\text{Co}_{0.2}\text{O}_2$ regression analysis.

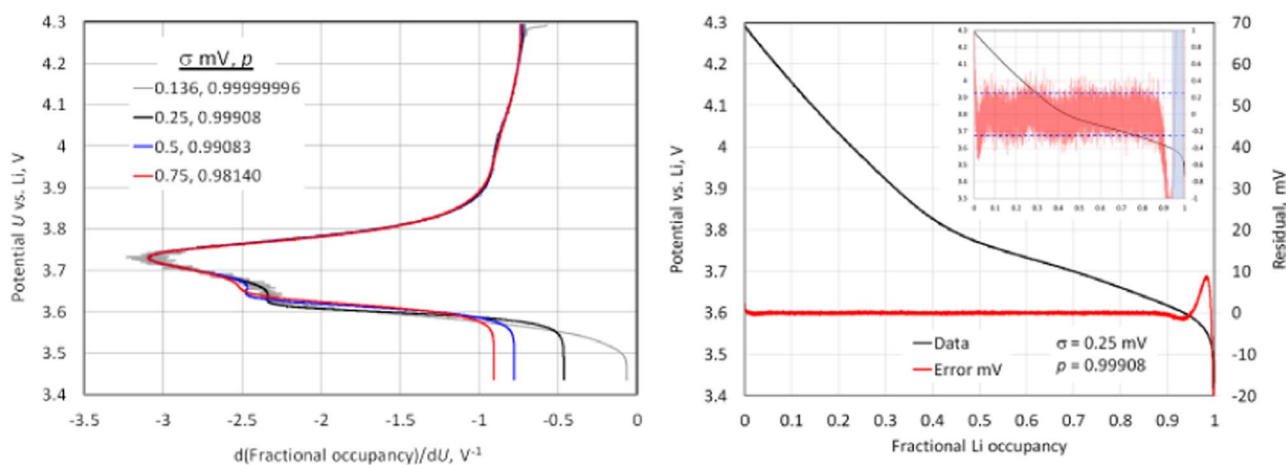


Figure 12. Left: regressed potentials for the $\text{Ni}_{0.6}\text{Mn}_{0.2}\text{Co}_{0.2}\text{O}_2$ for varying σ (and ρ) values. Right: Residual $U(x_i) - a_i$ in the smoothing spline fit shown in Figure 11 ($\sigma = 0.25$ mV). Within the inset figure, the dashed lines correspond to ± 0.25 mV, consistent with $\sigma = 0.25$ mV, and the shaded portion indicates the voltage range excluded (below 3.6 V) from use in calculating the squared residuals to determine ρ (cf. Eq. B23).

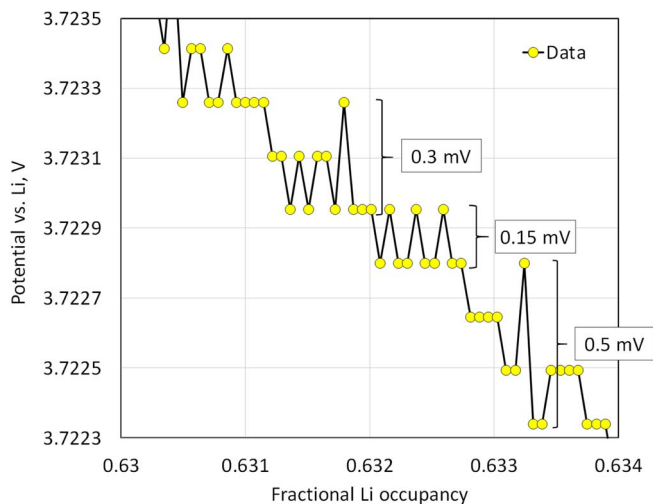


Figure 13. Noise in the $\text{Ni}_{0.6}\text{Mn}_{0.2}\text{Co}_{0.2}\text{O}_2$ experimental data near the peak in the dx/dU plot (cf. Figure 11). The noise is roughly twice that for the graphite trace (cf. Figure 4).

The unwanted noise in Figure 14 can be removed by going to the adaptive window width, corresponding to Algorithm 2 of Appendix A. Shown in Figure 15 are results for the adaptive-width windowed smoothing with a cubic polynomial. (The adaptive window regression yields results that are indistinguishable from those of the constant window regression for $U(x)$ and dx/dU as plotted in Figure 2 and Figure 5, respectively.) Larger window widths result about the larger potential plateaus, corresponding to larger reaction peaks, which gives more smoothing about the reaction peaks. Because the smoothing can be done locally with the adaptive window, the smaller value in the standard deviation in the data (0.15 mV, consistent with Figure 4) could be employed, allowing for more detail to be captured in the data without excessive noise in determining $p'' \approx d^2U/dx^2$, as is made clear by comparing and contrasting Figure 16 with Figure 14. Notice that in contrast to the fixed-width windowing and the smoothing spline algorithms, the adaptive-width approach does not require any restriction on the voltage range over which smoothing is determined, as the global error measures of Eqs. A6 or B8, for the fixed-width algorithm and the smoothing spline, respectively, are replaced by the local error measure of Eq. A9. Numerical values for the reaction characteristics extracted from the data using Algorithm 2 of Appendix A, without any other user judgement, are provided in Table I. In summary, the use of the adaptive window as described in Algorithm 2,

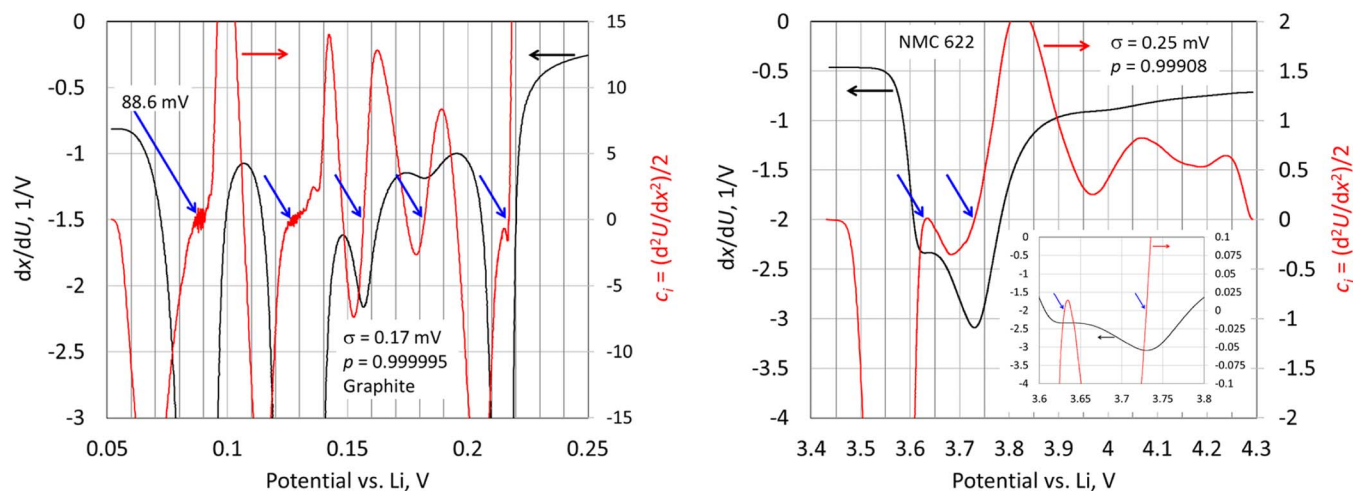


Figure 14. Smoothing spline analysis for reaction identification, characterized by a positive-slope crossing by d^2U/dx^2 about the line $d^2U/dx^2 = 0$, corresponding to a minimum (a reaction peak) in the dx/dU trace. Left plot: graphite. Right plot: NMC 622. Significant noise is seen for the two largest peaks in the case of graphite, corresponding to the regions identified by the two leftmost arrows. The inset plot for the NMC 622 analysis shows the positive-slope crossing by d^2U/dx^2 for the two reactions.

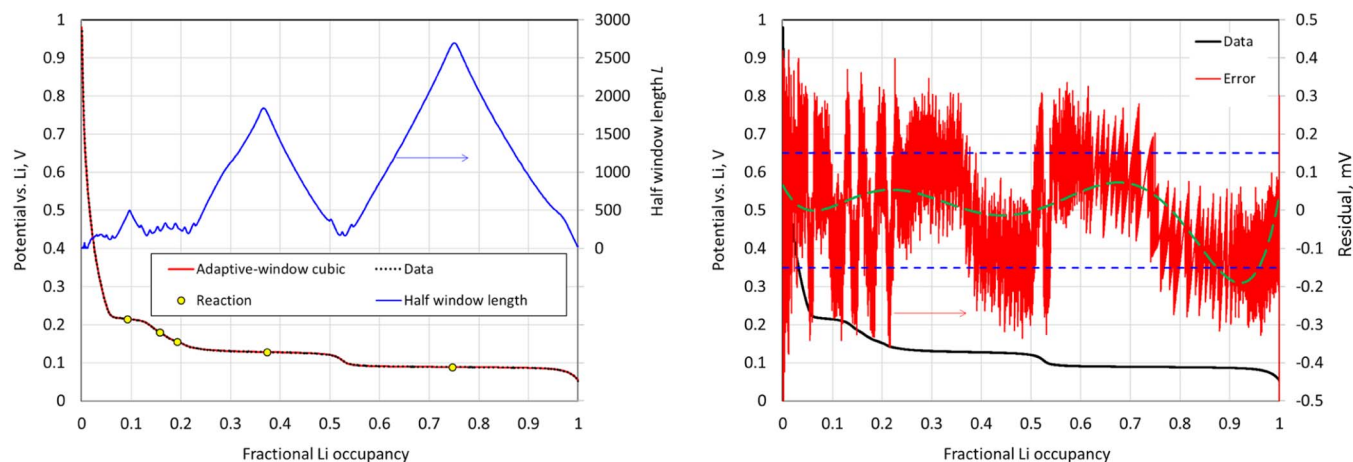


Figure 15. Adaptive-width windowed smoothing with a cubic polynomial (cf. Algorithm 2 of Appendix A). Left figure: lithiated graphite open-circuit potential and window half width L . The circles in the left plot correspond to the reaction potentials, as discussed in the context of Figure 16. The dashed lines correspond to ± 0.15 mV, consistent with $\sigma = 0.15$ mV (cf. Figure 4). The dashed green curve is a 6th order polynomial fit of the entire data set to represent the low-frequency trend in the data.

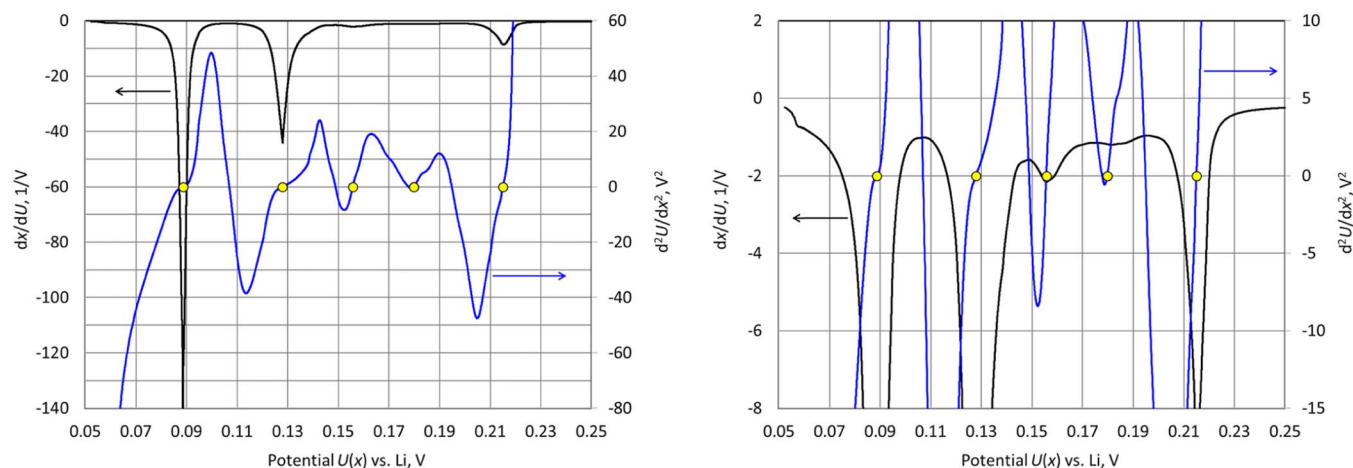


Figure 16. Adaptive-width windowed smoothing with a cubic polynomial for reaction identification, characterized by a positive-slope crossing by d^2U/dx^2 about the line $d^2U/dx^2 = 0$ (denoted by circles), corresponding to a minimum (a reaction peak) in the dx/dU trace for lithiated graphite. These results can be compared and contrasted with those of Figure 14. The right plot expands the ordinate values relative to those of the left plot.

Table I. Regressed values corresponding to the symbols (circles) in Figure 15 and Figure 16.

x	Smoothed $U(x)$, V	dU/dx , V	dx/dU , V ⁻¹
0.747	0.0886	-0.0071	-141
0.373	0.128	-0.0228	-43.8
0.192	0.156	-0.4646	-2.15
0.158	0.180	-0.842	-1.19
0.0918	0.215	-0.118	-8.51

while computationally more difficult, does allow us to provide an algorithm that satisfies the objective outlined in Figure 1; i.e., the autonomous identification of reactions by means of a nonparametric regression routine.

Summary

The focus of this work is enabling accurate DVS analyses, which requires an accurate method to differentiate numerically experimental data. Techniques employed for nonparametric regression and feature recognition are used to identify when electrochemical reactions take place in two commercially important intercalation electrode materials, lithiated graphite and nickel-manganese-cobalt oxide. The differentiation challenge is particularly significant for intercalation electrodes that undergo staging (similar to phase changes), as the peaks in plots of dx/dU vs. U are very sharp, as is seen for lithiated graphite (cf. Figure 5 and Figure 9). We analyze a moving window of data that are fit with a cubic polynomial (Appendix A) and a cubic smoothing spline (Appendix B). For the cubic polynomial, we investigate constant and adaptive window lengths. The three techniques are shown to work well for constructing DVS plots. For all three methods, we show how the standard deviation in the measurement method and the sum of the squared differences between the data and the regression fit can be used to estimate how large the moving window frame should be (Eqs. A6 and A9) and what the spline smoothing parameter should be (Eq. B23).

For the smoothing spline, we provide a rational means to nondimensionalize and scale the equation that is minimized (see Eqs. B9–B12); these matters appear to have been overlooked in the literature associated with smoothing splines. Allowing the window length for the cubic polynomial regression to adapt and vary yields the best results of all methods examined, as the adaptive window length enables more smoothing in the data where needed (e.g., around reaction peaks in dx/dU vs. U plots). The adaptive window method also removes the necessity for the user to restrict the range of data used in determining the smoothing.

An open issue that we do not address in this work is that of constraining the regression procedures to ensure that the fit function always has a negative slope, $dU/dx < 0$, as required by thermodynamic principles (i.e., the potential of a lithium-ion battery must always decline on discharge). Constrained regressions could be employed to enforce a negative slope, but the additional computational complexity and cost are significant. Mammen et al.³¹ discuss this topic in the general context of polynomial regression, kernel smoothing, and smoothing splines. Turlach³² discusses the topic in detail for smoothing splines. More research accommodating the methods discussed in Refs. 31 and 32 in the context of differential voltage spectroscopy may prove beneficial. In the work presented here, rather than imposing a constraint on the slope, we use any violation of that constraint as a signal that more smoothing is necessary.

Appendix A. Windowed Polynomial Smoothing

This Appendix describes methods for approximating a smooth function $f(x)$ given noisy data points

$$y_i = f(x_i) + \epsilon_i, \quad i = 1, \dots, N \quad [A1]$$

where ϵ_i , $i = 1, \dots, N$, are random errors. The methods will use polynomial approximations to subsets of the points to smooth out the random errors. We will begin by reviewing how to compute a least squares polynomial approximation to a set of points, and then describe how this can be used iteratively on a sliding window of data points to approximate a long sequence of points with a succession of low-order polynomials.

For integer $d \geq 1$, let

$$p(x) = p(x; \mathbf{a}) = p(x; a_0, \dots, a_d) = a_0 + a_1x + \dots + a_dx^d \quad [A2]$$

be a polynomial of degree d with coefficient array $\mathbf{a} = [a_0 \dots a_d]^T$. Given any $d + 1$ points (x_i, y_i) , for $i = 0, \dots, d$, all x_i distinct, there exists a unique polynomial $p(x)$ that interpolates the points. Due to effects related to the Runge phenomenon,^b even for noise-free data points, this interpolating polynomial may be a poor approximation to $f(x)$, and especially as N grows, $p(x)$ may oscillate wildly. The problem of oscillations is exacerbated by noise and further amplified if one evaluates derivatives of the polynomial.

A remedy for this situation is to use a polynomial of lower degree to approximate the data points, and possibly to use different low-order polynomials on different subsets of the points. As one option, one might opt to partition the data into intervals and construct a piecewise polynomial approximation with continuity conditions enforced where the intervals meet; spline functions are constructions of this type. In this section, we consider instead an approach that, for each abscissa x_i , approximates $f(x)$ at $x = x_i$ using a polynomial fit to the points within a sliding window containing x_i . For interior points, this window will be symmetrically centered on x_i , but for points near either extreme of the range of the data, the window must necessarily become asymmetric.

To describe this method, we begin by considering how to fit a polynomial to n points. Later we will address the problem of selecting for each abscissa x_i which subset of points to use in constructing an approximation there.

For $n > d$ distinct points, we may find the unique least squares approximant, that is, the polynomial $p(x; \mathbf{a})$ that minimizes the sum of squares

$$S(\mathbf{a}) = \sum_{i=1}^n (p(x_i; \mathbf{a}) - y_i)^2$$

One may consider the optimal $p(x; \mathbf{a})$ to approximate the first $d + 1$ terms of the Taylor series for $f(x)$ centered at $x = 0$. The differences $p(x_i; \mathbf{a}) - y_i$ are called the *residuals* of the fit.

Computation of the least squares polynomial is most easily described in matrix notation. First, we gather the data into arrays $\mathbf{x} = [x_1 \dots x_n]^T$ and $\mathbf{y} = [y_1 \dots y_n]^T$ and let

$$p(\mathbf{x}; \mathbf{a}) = [p(x_1; \mathbf{a}) \dots p(x_n; \mathbf{a})]^T$$

Then, the objective function $S(\mathbf{a})$ can be rewritten as $S(\mathbf{a}) = \|p(\mathbf{x}; \mathbf{a}) - \mathbf{y}\|^2$. It is convenient to explicitly show the linear dependence of $p(\mathbf{x}; \mathbf{a})$ on its coefficients via the so-called the Vandermonde matrix,^{23,28} the $n \times (d + 1)$ matrix with entries $V_{ij} = x_i^{j-1}$:

$$V = \begin{bmatrix} 1 & x_1 & \dots & x_1^d \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^d \end{bmatrix}$$

This allows us to write

$$p(\mathbf{x}; \mathbf{a}) = V\mathbf{a}, \quad S(\mathbf{a}) = (V\mathbf{a} - \mathbf{y})^T (V\mathbf{a} - \mathbf{y})$$

Equating the partial derivatives of S with respect to \mathbf{a} to zero, we obtain the following optimality criterion, often called the *normal equation* for least squares:²⁵

$$V^T V \mathbf{a} = V^T \mathbf{y}$$

For $d < n$ with at least $(d + 1)$ of the x_i distinct, it can be shown that square matrix $V^T V$ is nonsingular,²³ hence there is a unique set of least squares coefficients:

$$\mathbf{a} = (V^T V)^{-1} V^T \mathbf{y} \quad [A3]$$

In more detail, writing the normal equation as $M\mathbf{a} = \mathbf{b}$, with $M = V^T V$, $\mathbf{b} = V^T \mathbf{y}$, we have formulas for the entries of M and \mathbf{b} as

$$M_{ij} = \sum_{k=1}^n x_k^{i+j-2} \quad \text{and} \quad \mathbf{b}_i = \sum_{k=1}^n x_k^{i-1} y_k \quad [A4]$$

The coefficients of the least squares fit are $\mathbf{a} = M^{-1}\mathbf{b}$. As detailed in Ref. 25 the same least-squares solution can be computed more robustly using either the pseudoinverse V^\dagger of V to write $\mathbf{a} = V^\dagger \mathbf{y}$, or a so-called QR decomposition of $V = QR$, where Q is orthogonal and R is upper triangular, in which case $\mathbf{a} = R^{-1}Q^T \mathbf{y}$. The algorithms for pseudoinversion and QR decomposition are available in any modern software package for linear algebra, so we omit further details here.

^bSee de Boor¹⁹ and especially Berrut and Trefethen²⁴ for discussions of the Runge phenomenon. The phenomenon can be suppressed by careful spacing of the samples near the ends of the interval, such as in Chebyshev interpolation, but we seek a method not dependent on such spacing. Indeed, experimental data is often produced at uniform intervals for which the Runge phenomenon is severe.

We note that for numerical robustness, it is also a good idea to scale and translate the x -coordinates so that the transformed coordinates range over 0 to 1 or -1 to 1.

Special case: uniformly spaced points.—Suppose that the x -coordinates of samples are uniformly spaced, that is, $x_i = ih$ for some constant increment h . Suppose further that we elect to fit a polynomial to $n = 2L + 1$ consecutive points, L integer. Then, if we translate coordinates to place the origin at the middle point and re-index the points as $x_{-L}, x_{-L+1}, \dots, x_{L-1}, x_L$, we have a symmetrical arrangement where $x_{-k} = -x_k$, $k = 1, \dots, L$. As a consequence, by Eq. A4, for $i + j$ equal to any odd number, $M_{ij} = 0$. To take advantage of this, we rearrange the ordering of the coefficients in a to place the even terms first and followed by the odd ones; for example, for $d = 3$ (cubic polynomial), we use $\mathbf{a} = [a_0 \ a_2 \ a_1 \ a_3]^T$. When the columns of V are rearranged to match, the matrix M becomes block diagonal:

$$M = V^T V = \begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix}$$

Rather than writing general expressions for M_1 and M_2 , we illustrate for the specific case of cubic polynomials, for which M and the associated right-hand vector \mathbf{b} become

$$M = \begin{bmatrix} \sum x_i^2 & \sum x_i^4 & 0 & 0 \\ \sum x_i^4 & \sum x_i^6 & 0 & 0 \\ 0 & 0 & \sum x_i^2 & \sum x_i^4 \\ 0 & 0 & \sum x_i^4 & \sum x_i^6 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} \sum y_i \\ \sum x_i^2 y_i \\ \sum x_i y_i \\ \sum x_i^3 y_i \end{bmatrix}$$

where each sum is taken over $i = -L, \dots, L$. The block diagonal structure allows one to solve for the odd and even coefficients independently, thereby reducing the original 4×4 matrix solve to two 2×2 solves. The generalization for larger d is straightforward. Moreover, one may simplify the entries of M using $\sum x_i^k = h^k \sum i^k$. Because we invoked a translation of the origin to x_{mid} , in the original coordinates the polynomial becomes $p(x - x_{mid}; \mathbf{a})$.

Moving window smoothing.—It often happens that a single low-order polynomial cannot adequately approximate a given data set. For example, a polynomial of degree d has at most $d - 1$ stationary points (maxima, minima, and inflections) and the function to be approximated could have more than this. One approach is to break the data set into pieces and fit a polynomial on each piece separately. Enforcing continuity of the piecewise polynomial and some number of derivatives yields a spline function. The smoothing spline discussed in the next section falls into this class.

Moving window smoothing takes a related, but different, approach. For each abscissa x_i , we propose to compute a polynomial $p_{k_1:k_N}(x)$ that is the least squares fit to a window of neighboring data points (x_k, y_k) , $k = k_1, \dots, k_N$, choosing k_1 and k_N to include i , i.e., $k_1 \leq i \leq k_N$. Then, evaluating this polynomial at x_i gives a value $f_i = p_{k_1:k_N}(x_i)$ that we take as the approximation to $f(x_i)$. We may similarly use the derivatives of $p_{k_1:k_N}(x)$ evaluated at x_i to approximate the derivatives of $f(x)$ at x_i .

One may develop a variety of smoothing algorithms following this same basic plan. The choices that must be made are the degree d of the polynomials and the window to use around each point. We have found that cubic polynomials ($d = 3$) work well in our applications. In what follows, we propose two different prescriptions for choosing windows around the points. The first is simpler to implement, while the second shows clear advantage on some more challenging data sets.

Algorithm 1: constant-width windowed smoothing.—For this approach, we use an odd number, $n = 2L + 1$, of successive points for each approximation. Assume the given data points are numbered from 1 to N . For points in the middle of the data set, specifically, for indexes $L < i < N - L$, we use a window centered on x_i to approximate there, while the ends of the data set must use unbalanced windows. Specifically, we use

$$f_i(L) = \begin{cases} p_{1:(2L+1)}(x_i), & i \leq L; \\ p_{(i-L):(i+L)}(x_i), & L < i < N - L; \\ p_{(N-2L):N}(x_i), & i \geq N - L. \end{cases} \quad \text{[A5]}$$

This leaves just the window half-width, L , to be specified. We will determine L so as to make residuals in the fit consistent with our knowledge of the accuracy of the process that produced the data. Suppose that the errors in the data points are independent, zero-mean, with a standard deviation of σ , that is, the errors ϵ_i in Eq. A1 are independent with expectations $E(\epsilon_i) = 0$ and $E(\epsilon_i^2) = \sigma^2$. For a given L , we obtain a sum of squared residuals^c

$$SSR(L) = \sum_{i=1}^N (f_i(L) - y_i)^2. \quad \text{[A6]}$$

We expect $SSR \approx N\sigma^2$. (For more on squared residuals and nonparametric fitting, see chapter 2 of Eubank.¹¹) If $SSR(N/2) \leq N\sigma^2$, that is, if our polynomial fit to the entire data set falls within the expected residual, then we accept that polynomial as the smoothed fit. Otherwise, some smaller width is appropriate. For $2L \leq d$, the polynomials in Eq. A5

^cIn the statistics literature, instead of SSR, the term *residual sum of squares* and abbreviation RSS are sometimes used to mean the same thing.

interpolate the data, thus giving $SSR(L \leq d/2) = 0$. As L increases from its minimum size to its maximum size, at some point $SSR(L)$ must cross its target value of $N\sigma^2$, and we aim to choose the largest L such that $SSR(L - 1) \leq N\sigma^2$ as the window width for the smoothed approximation.

To determine the appropriate value of L , it is not necessary to increment through all its possible values. Instead, one may approach the problem using any available numerical root finder, such as bisection or the secant method, to solve $SSR(L) = N\sigma^2$. Of course, the iterates for L must be rounded off to integers, and the root finder must terminate upon finding an L such that

$$SSR(L - 1) \leq N\sigma^2 \leq SSR(L). \quad \text{[A7]}$$

If $SSR(N/2) > N\sigma^2$, we know that such an L must exist, although it may not be unique. Ideally, if there is more than one such width, we would use the largest one, but our experience is that any L satisfying this criterion is acceptable.

It is not necessary to pinpoint L meeting the criterion in Eq. A7, as a nearby value will give almost the same fit. According to Reinsch's recommendation for the smoothing spline,^{14,19} one could terminate the search upon finding a window width such that

$$\left| \frac{SSR(L)}{\sigma^2} - N \right| \leq \sqrt{2N}. \quad \text{[A8]}$$

This criterion is justified on the assumption that the errors ϵ_i are independent and Gaussian, in which case the sum of squared errors is a chi-square distribution that has a mean of $N\sigma^2$ and a standard deviation of $\sqrt{2N}\sigma^2$. More precisely, since we are solving for $(d + 1)$ coefficients in each polynomial fit, it may be more appropriate to seek a value of L such that $SSR(L) = (N - d - 1)\sigma^2$, but in our work $N \gg d$, so this minor adjustment would be of no substantial consequence.

Immediately below, we provide a pseudo-code summary of Algorithm 1.

Begin Constant-width Window Smoothing Algorithm

Inputs: data arrays \mathbf{x} , \mathbf{y} of length N ; standard deviation σ .

Outputs: half-width L ; smoothed function values \mathbf{f} .

Comment: floor(z) = greatest integer less than z

Initialize: $L_0 = \text{floor}(d/2)$, $S_0 = 0$; $L_1 = \text{floor}(N/2)$;

Evaluate: $[S_1, \mathbf{f}] = SSR(\mathbf{x}, \mathbf{y}, L_1)$;

Comment: check if one polynomial fits the whole data set

If $S_1 \leq N\sigma^2$, $L = L_1$; Return; end.

Comment: use bisection to find window width that satisfies Eq. A7

While $L_1 > L_0 + 1$

$L = \text{nearest_integer}((L_0 + L_1)/2)$

Evaluate $[S_{test}, \mathbf{f}_{test}] = SSR(\mathbf{x}, \mathbf{y}, L)$

If $S_{test} \leq N\sigma^2$, then $L_0 = L$, else $L_1 = L$, $\mathbf{f} = \mathbf{f}_{test}$.

$L = L_1$; Return.

End Algorithm

Algorithm 2: adaptive-width windowed smoothing.—Some functions, including the open-circuit voltage curve for graphite, have regions with very low curvature punctuated by rapid transitions of high curvature. For such a function, the flat regions can be well approximated with a wider window than can be used where curvature changes quickly, so it is inappropriate to use a fixed window width. Letting the window width vary and denoting its value at point i as L_i , we may modify Eq. A5 to use $f_i(L_i)$ instead of a fixed L . More precisely, we will determine an initial window of width L_1 that is used for all f_i , $i \leq L_1 + 1$, and a final window of width L_N that is used for all f_i , $i \geq N - L_N$. In between these, the L_i must respect the limits of the indexes, $L_i < i$ and $L_i < N - i$, and also the lower limit, $L_i \geq d/2$, below which the local polynomial is an exact interpolant of all points in the window. For the sake of robustness, it can be useful to set a somewhat larger minimum width, and in fact, we used cubic polynomials, $d = 3$, and set $L_i \geq 6$ in this article.

To determine the window widths, let us define a local sum of squared residuals for the window centered on point i , named SSR_i to distinguish it from SSR above:

$$SSR_i(i, L) = \sum_{k=i-L}^{i+L} (p_{(i-L):(i+L)}(x_k) - y_k)^2. \quad \text{[A9]}$$

This formula is only defined for $i > L$ and $i \leq N - L$, otherwise the indexes in the fit extend beyond the ends of the data set. As we did in Eq. A5 for the fixed window width algorithm, we will have to use an asymmetric set of points to approximate function values near the beginning and end of the data set.

As in the fixed-width case, we seek a window width compatible with the assumed noise covariance, that is, we seek a width L_i such that $SSR_i(i, L_i) = (2L_i + 1)\sigma^2$ for all i . Since L_i is an integer, we cannot generally obtain equality, so instead we seek a window width that simultaneously satisfies

$$SSR_i(i, L_i - 1) < (2L_i - 1)\sigma^2 \quad \text{and} \quad SSR_i(i, L_i) \geq (2L_i + 1)\sigma^2. \quad \text{[A10]}$$

We expect the window widths for neighboring points to be close to each other. Accordingly, after finding the initial width for the head of the data set, for every point thereafter, we use L_i as the initial guess for the search for L_{i+1} . Since the window width only changes

allowed, and we obtain a least-squares linear regression of the data. There is a problem with this well-used formulation, however: the first term on the right side of Eq. B8 is dimensionless, whereas the second term has dimensions of $(u_f)^2/(u_x)^3$, where u_f are the units of $f(x)$ and u_x are the units of x . In this work, x is dimensionless, but to fix Eq. B8 so that it has a more general applicability, we replace it with

$$P = \frac{p}{N} \sum_{i=1}^N \left(\frac{y_i - a_i}{\sigma_i} \right)^2 + (1-p) \frac{(x_N - x_1)^3}{s^2} \int_{x_1}^{x_N} [f''(x)]^2 dx \tag{B9}$$

where both terms on the right side are dimensionless, even if x and y have dimensions associated with them, and we suggest employing

$$s = \sqrt{\frac{1}{N} \sum_{i=1}^N [y_i - l(x_i)]^2} \tag{B10}$$

that is, s the square root of the sum of the squared deviation of the data $y_i(x_i)$ from a best-fit line $l(x_i)$ for the data, which, being a line, has no curvature, thus providing a reference to scale $f''(x)$. The units of s are those of the data y_i and the function $f(x)$. One can argue as to whether s formulated in Eq. B10 is the right scaling for $f''(x)$, and we shall revisit this topic in the Results and Discussion section. Equation B9 has an additional benefit insofar as one can then employ an ascending or descending series in the independent variable, as the sign $(x_N - x_1)^3 = \text{sign}(dx)$, where dx is the differential increment in the integral of Eq. B9. (For the treatments in Refs. 19 and 26 the independent variable must be in ascending order. This is not a problem for problems having ascending time as the independent variable.) It is common to assume a constant standard deviation in the data, $\sigma_i = \sigma$, an assumption we employ in this work, and this allows us to recast Eq. B9 with the lone dimensionless group λ :

$$P_\lambda = \sum_{i=1}^N (y_i - a_i)^2 + \lambda \int_{x_1}^{x_N} [f''(x)]^2 dx \tag{B11}$$

where

$$\lambda = N(x_N - x_1)^3 \frac{\sigma^2}{s^2} \left(\frac{1-p}{p} \right) \tag{B12}$$

Solving the matrix equation for the smoothing spline coefficients.—In discretized form,^d Eq. B11 can be written as

$$P_\lambda = \sum_{i=1}^N (y_i - a_i)^2 + \frac{4}{3} \lambda \sum_{i=1}^{N-1} h_i (c_i^2 + c_i c_{i+1} + c_{i+1}^2) = (\mathbf{y} - \mathbf{a})^T (\mathbf{y} - \mathbf{a}) + \frac{2}{3} \lambda \mathbf{c}^T \mathbf{R} \mathbf{c} \tag{B13}$$

Differentiating P_λ with respect to \mathbf{a} and setting the result equal to zero leads to

$$\mathbf{a} = \mathbf{y} - 2\lambda \mathbf{Q} \mathbf{c} \tag{B14}$$

Eliminating \mathbf{a} from Eq. B14 by means of Eq. B7 leads to

$$[\mathbf{R} + 6\lambda \mathbf{Q}^T \mathbf{Q}] \mathbf{c} = 3\mathbf{Q}^T \mathbf{y} \tag{B15}$$

where $\mathbf{Q}^T \mathbf{Q}$ is an $(N-2) \times (N-2)$ symmetric, pentadiagonal matrix. For notational convenience, we seek a solution to

$$\mathbf{A} \mathbf{c} = \boldsymbol{\beta} \text{ where } \mathbf{A} = \mathbf{R} + 6\lambda \mathbf{Q}^T \mathbf{Q} \text{ and } \boldsymbol{\beta} = 3\mathbf{Q}^T \mathbf{y} \tag{B16}$$

In expanded form, for $N = 7$, \mathbf{A} can be written as the matrix sum

$$\mathbf{A} = \mathbf{R} + 6\lambda \mathbf{Q}^T \mathbf{Q} = \begin{bmatrix} 2(h_1 + h_2) & h_2 & & & & & \\ h_2 & 2(h_2 + h_3) & h_3 & & & & \\ & h_3 & 2(h_3 + h_4) & h_4 & & & \\ & & h_4 & 2(h_4 + h_5) & h_5 & & \\ & & & h_5 & 2(h_5 + h_6) & & \\ & & & & & & \end{bmatrix} + 6\lambda \begin{bmatrix} \eta_1^2 + \eta_2^2 + \eta_3^2 & -\eta_{12}\eta_2 - \eta_2\eta_{23} & \eta_2\eta_3 & & & & \\ -\eta_{12}\eta_2 - \eta_2\eta_{23} & \eta_2^2 + \eta_{23}^2 + \eta_3^2 & -\eta_{23}\eta_3 - \eta_3\eta_{34} & \eta_3\eta_4 & & & \\ \eta_3\eta_2 & -\eta_{23}\eta_3 - \eta_3\eta_{34} & \eta_3^2 + \eta_{34}^2 + \eta_4^2 & -\eta_{34}\eta_4 - \eta_4\eta_{45} & \eta_4\eta_5 & & \\ & \eta_3\eta_4 & -\eta_{34}\eta_4 - \eta_4\eta_{45} & \eta_4^2 + \eta_{45}^2 + \eta_5^2 & -\eta_{45}\eta_5 - \eta_5\eta_{56} & & \\ & & \eta_4\eta_5 & -\eta_{45}\eta_5 - \eta_5\eta_{56} & \eta_5^2 + \eta_{56}^2 + \eta_6^2 & & \end{bmatrix} \tag{B17}$$

^dHere we employ the fact that for any straight line $w(x)$, $\int_{x_i}^{x_{i+1}} [w(x)]^2 dx = \left(\frac{x_{i+1} - x_i}{3} \right) [w(x_i)^2 + w(x_i)w(x_{i+1}) + w(x_{i+1})^2]$, along with Eq. B2, $S_i''(x_i) = 2c_i$.

and $\boldsymbol{\beta}$ can be written as

$$\boldsymbol{\beta} = 3\mathbf{Q}^T \mathbf{y} = 3 \begin{bmatrix} \eta_1 & -\eta_{12} & \eta_2 & & & & \\ & \eta_2 & -\eta_{23} & \eta_3 & & & \\ & & \eta_3 & -\eta_{34} & \eta_4 & & \\ & & & \eta_4 & -\eta_{45} & \eta_5 & \\ & & & & \eta_5 & -\eta_{56} & \eta_6 \\ & & & & & & \eta_7 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{bmatrix} = 3 \begin{bmatrix} \eta_1 y_1 - \eta_{12} y_2 + \eta_2 y_3 \\ \eta_2 y_2 - \eta_{23} y_3 + \eta_3 y_4 \\ \eta_3 y_3 - \eta_{34} y_4 + \eta_4 y_5 \\ \eta_4 y_4 - \eta_{45} y_5 + \eta_5 y_6 \\ \eta_5 y_5 - \eta_{56} y_6 + \eta_6 y_7 \end{bmatrix} \tag{B18}$$

More generally, the symmetric, pentadiagonal matrix \mathbf{A} has elements

$$\begin{aligned} A_{i,i} &= 2(h_i + h_{i+1}) + 6\lambda(\eta_i^2 + \eta_{i,i+1}^2 + \eta_{i+1}^2), \quad i = 1 \text{ to } N - 2 \\ A_{i,i+1} &= A_{i+1,i} = h_{i+1} - 6\lambda(\eta_{i,i+1}\eta_{i+1} + \eta_{i+1}\eta_{i+1,i+2}), \quad i = 1 \text{ to } N - 3 \\ A_{i,i+2} &= A_{i+2,i} = 6\lambda\eta_{i+1}\eta_{i+2}, \quad i = 1 \text{ to } N - 4 \end{aligned} \tag{B19}$$

and the vector $\boldsymbol{\beta}$ has elements

$$\beta_i = 3(\eta_i y_i - \eta_{i,i+1} y_{i+1} + \eta_{i+1} y_{i+2}), \quad i = 1 \text{ to } N - 2 \tag{B20}$$

Equation B16, $\mathbf{A} \mathbf{c} = \boldsymbol{\beta}$, can be solved by using efficient methods for pentadiagonal systems. From a storage perspective, \mathbf{A} and $\boldsymbol{\beta}$ can be represented as

$A_{i,i}$	$A_{i,i+1}$	$A_{i,i+2}$	$\boldsymbol{\beta}$	\mathbf{c}
A_{11}	A_{12}	A_{13}	β_1	c_2
A_{22}	A_{23}	$A_{2,4}$	β_2	c_3
\vdots	\vdots	\vdots	\vdots	\vdots
\vdots	\vdots	$A_{N-4,N-2}$	\vdots	\vdots
\vdots	\vdots	\vdots	\vdots	\vdots
\vdots	$A_{N-3,N-2}$	\vdots	\vdots	\vdots
$A_{N-2,N-2}$	\vdots	\vdots	β_{N-2}	c_{N-1}

No entries are shown in B21 for $A_{i+1,i}$ or $A_{i+2,i}$, as \mathbf{A} is symmetric (hence, $A_{i,i+1} = A_{i+1,i}$ and $A_{i,i+2} = A_{i+2,i}$).

In this work, $\mathbf{A} \mathbf{c} = \boldsymbol{\beta}$ (reflected by entries of B21) was solved using the algorithm described in Cheney and Kincaid.²⁷ Upon solving Eq. B16, one can use Eq. B14 to recover \mathbf{a} :

$$\begin{aligned} \text{For } i = 1 & \\ a_1 &= y_1 - 2\lambda\eta_1 c_2 \\ \text{For } i = 2 \text{ to } N - 1 & \\ a_i &= y_i - 2\lambda(\eta_{i-1} c_{i-1} - \eta_{i-1,i} c_i + \eta_i c_{i+1}) \\ \text{For } i = N & \\ a_N &= y_N - 2\lambda\eta_{N-1} c_{N-1} \end{aligned} \tag{B22}$$

Having \mathbf{a} and \mathbf{c} , one can employ Eqs. B4 and B5 to determine the d_i and b_i coefficients, respectively.

Determination of the smoothing parameter p .—The same logic used to determine an appropriate value of L for polynomial regression can be applied to determine an appropriate value of p . Hence, an equation for the smoothing spline that is analogous to that of Eq. A9, with $SSR = N\sigma^2$, is employed:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N [f(x_i) - S(x_i; a_i, b_i, c_i, d_i, p)]^2} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - a_i)^2} \tag{B23}$$

This provides a (nonlinear) relation that can be used to determine an estimate for λ , and thus p , (cf. Eq. B12) once the standard deviation in the measurement σ is provided. Helpful expositions on Generalized Cross Validation devoted to smoothing splines can be found in Refs. 29 and 30; and the use of Eq. B23 represents a pragmatic method for addressing a complicated subject.

ORCID

Mark W. Verbrugge <https://orcid.org/0000-0001-6100-3238>
 Charles W. Wampler <https://orcid.org/0000-0002-9790-1600>
 Daniel R. Baker <https://orcid.org/0000-0002-7642-4067>

References

1. C. Pastor-Fernández, K. Uddin, G. H. Chouchelamane, W. D. Widanage, and J. Marco, *J. Power Sources*, **360**, 301 (2017).
2. R. Deshpande, M. W. Verbrugge, Y.-T. Cheng, J. Wang, and P. Liu, *J. Electrochem. Soc.*, **159**, A1730 (2012).

3. Y. Li, M. Abdel-Monema, R. Gopalakrishnan, M. Bercibar, E. Nanini-Maury, N. Omar, P. van den Bossche, and J. Van Mierlo, *J. Power Sources*, **373**, 40 (2018).
4. L. Zheng, J. Zhu, D. D.-C. Lu, G. Wang, and T. He, *Energy*, **150**, 759 (2018).
5. J. Newman and K. E. Thomas-Alyea, *Electrochemical Systems*, 3rd Edition, Wiley Interscience, 2004.
6. K. E. Thomas, R. M. Darling, and J. Newman, "Mathematical Modeling of Lithium Batteries," in: *Advances in Lithium-Ion Batteries*, W. van Schalkwijk and B. Scrosati, (eds.), Kluwer, Dordrecht (2002), Chap. 12.
7. M. W. Verbrugge and B. J. Koch, *J. Electrochem. Soc.*, **143**, 24 (1996).
8. D. R. Baker and M. W. Verbrugge, *J. Electrochem. Soc.*, **159**, A1341 (2012).
9. A. J. Bard and L. R. Faulkner, *Electrochemical Methods, Fundamentals and Applications*, 2nd edition, Wiley, New York, NY 2001. See section 6.6 for cyclic voltammetry, current peaks, and multi-reaction systems.
10. M. W. Verbrugge, D. R. Baker, B. J. Koch, X. Xiao, and W. Gu, *J. Electrochem. Soc.*, **164**, E3243 (2017).
11. R. L. Eubank, *Nonparametric Regression and Spline Smoothing*, 2nd edition, CRC Press, 1999.
12. D. R. Baker and M. W. Verbrugge, *J. Electrochem. Soc.*, **165** (16) A1 (2018).
13. V. Ramadesigan, P. W. C. Northrop, S. De, S. Santhanagopalan, R. D. Braatz, and V. R. Subramaniana, *J. Electrochem. Soc.*, **159**, R31 (2012).
14. C. H. Reinsch, *Numer. Math.*, **10**, 177 (1967).
15. S. D. Conte and C. de Boor, *Elementary Numerical Analysis*, third edition, McGraw-Hill, 1980.
16. J. Richardson, P. Reiner, and B. M. Wilamowski, *Proc. IEEE INDIN 2015*, pp. 110.
17. G. Feng, *IEEE Transactions on Signal Processing*, **46**, 2790 (1998).
18. S. J. Russell and P. Norvig, *Artificial Intelligence. A Modern Approach*, third edition, Prentice Hall, 2010. See chapter 18 for agent (machine) learning and nonparametric regression.
19. C. de Boor, *A Practical Guide to Splines*, revised edition, Springer, 2001. See chapter 14 for smoothing splines.
20. P. Arora, M. Doyle, and R. E. White, *J. Electrochem. Soc.*, **146**, 3543 (1999).
21. R. D. Perkins, A. V. Randall, X. Zhang, and G. L. Plett, *J. Power Sources*, **209**, 318 (2012).
22. J. Cannarella and C. B. Arnold, *J. Electrochem. Soc.*, **162**, A1365 (2015).
23. O. Bretscher, *Linear Algebra with Applications*, 5th ed., Pearson, 2013.
24. J.-P. Berrut and L. N. Trefethen, *SIAM Review*, **46**, 501 (2004).
25. C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*, Prentice Hall, Englewood Cliffs, NJ, 1974.
26. D. S. G. Pollock, *A Handbook of Time-Series Analysis, Signal Processing and Dynamics*, Academic Press, 1999. See chapter 11 for smoothing splines.
27. W. Cheney and D. Kincaid, *Numerical Mathematics and Computing*, 6th edition, Thomson Brooks/Cole, 2008. See page 284 for the pentadiagonal solver.
28. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, 1989.
29. P. Craven and G. Wahba, *Numer. Math.*, **31**, 377 (1979).
30. Y. Lin and H. H. Zhang, *The Annals of Statistics*, **34**, 2272 (2006).
31. E. Mammen, J. S. Marron, B. A. Turlach, and M. P. Wand, *Statistical Science*, **16**, 232 (2001).
32. B. A. Turlach, *Computational Statistics*, **20**, 81 (2005).