energies

MDPI

Article

Incremental Capacity Analysis on Commercial Lithium-Ion Batteries using Support Vector Regression: A Parametric Study

Xuning Feng [1], Caihao Weng [2], Xiangming He [1,*], Li Wang [1], Dongsheng Ren [3], Languang Lu [3], Xuebing Han [3] and Minggao Ouyang [3,*]

1 Institute of Nuclear and New Energy Technology, Tsinghua University, Beijing 100084, China; fxn17@tsinghua.edu.cn (X.F.); wang-l@tsinghua.edu.cn (L.W.)
2 Naval Architecture and Marine Engineering, University of Michigan, Ann Arbor, MI 48104, USA; chsweng@gmail.com
3 State Key Laboratory of Automotive Safety and Energy, Tsinghua University, Beijing 100084, China; rds14@mails.tsinghua.edu.cn (D.R.); lulg@tsinghua.edu.cn (L.L.); hanxuebing@mail.tsinghua.edu.cn (X.H.)
* Correspondence: hexm@tsinghua.edu.cn (X.H.); ouymg@tsinghua.edu.cn (M.O.); Tel.: +86-010-6279-4226 (X.H.)

Received: 8 August 2018; Accepted: 28 August 2018; Published: 3 September 2018

Abstract: Incremental capacity analysis (ICA) has been used pervasively to characterize the degradation mechanisms of the lithium-ion batteries, and several online state-of-health estimation algorithms are built based on ICA. However, the stairs and the noises in the discrete sampled voltage data obstruct the calculation of the capacity differentiation over voltage ($\mathrm{d}Q/\mathrm{d}V$), therefore we need methods to fit the sampled voltage first. In this paper, the support vector regression (SVR) algorithm is used to smooth the sampled voltage curve using Gaussian kernels. A parametric study has been conducted to show how to enhance the performances of the SVR algorithm, including (1) speeding up the algorithm by downsampling; (2) avoiding overfitting and under-fitting using proper standard deviation σ in the Gaussian kernel; (3) making precise capture of the characteristic peaks. A novel method using linear approximation has been proposed to help judge the accuracy of the SVR algorithm in tracking the ICA peaks. And advanced SVR algorithms using double σ and using cost function that directly regulates the differentiation result have been proposed. The advanced SVR algorithm can make accurate curve fitting for ICA with overall error less than 1% (maximum 3%) throughout cycle lives, for four kinds of commercial lithium-ion batteries with $LiFePO_4$ and $LiNi_xCo_yMn_zO_2$ cathodes, making it promising to be further applied in online SOH estimation algorithms.

Keywords: lithium-ion batteries; state-of-health; incremental capacity analysis; support vector regression; curve fitting; energy storage

1. Introduction

The market of electric vehicles is growing at an unprecedented rate in recent years, driven by the motivations to find cleaner energy powertrain systems for future transportations [1–3]. Battery energy storage system, which the heart of the energy powertrain systems of electric vehicles, must have considerable energy density and long cycle life to guarantee the desired performances of electric vehicles. Lithium-ion batteries stand out from other kinds of energy storage technologies, because of their better energy efficiency and longer cycle life [4–7].

One challenge for the application of lithium-ion batteries in electric vehicles is the state-of-health (SOH) monitoring [8,9]. The SOH is 'a state' that reflects the aging condition in comparison with a fresh battery. The quantitative definition of SOH can be founded on the battery capacity [10,11] or the impedance [12–14]. The SOH monitoring is regarded as one of the key problems for battery management, because the SOH affects the accurate estimation of battery states, thereby determining the capability of power output of the battery system [15–17].

Incremental capacity analysis (ICA) is a method that is being pervasively used in battery SOH diagnosis in recent years, after its first proposal by Balewski & Brenet [18] and propagated by Dubarry et al. [19,20]. The ICA technique is performed by differentiating the capacity Q with respect to the voltage V, therefore the voltage plateaus that reflects the thermodynamics of battery will be transformed into $\mathrm{d}Q/\mathrm{d}V$ peaks, which are convenient for degradation analysis [21,22]. $\mathrm{d}Q/\mathrm{d}V$ is also called the incremental capacity (IC). The ICA is capable of revealing the mechanism of both the capacity loss and the impedance increment [23–25], and can be further applied for online SOH monitoring [26,27].

The major difficulty in performing the ICA comes from processing the differentiations of the discretized voltage data with sampling levels and noise. The sampling levels and noises can lead to large fluctuations in the IC curves, since differentiating discretized voltage at plateaus will return values near zero and infinity [28]. Curve fitting seems to be a must to perform ICA. A typical curve fitting method is to find continuous functions that can fit the discretized voltage data with minimum error. Then the differentiations can be acquired by differentiating the fitted continuous functions. There several available curve fitting methods [29–31], among which the SVR method is a kind of technique that comes from modern theory of machine learning (ML). Usually the ML-based methods are used for predicting the capacity loss [32–34], rather than fitting the curves of battery voltage. Weng et al. [35] firstly used the SVR to fit the battery voltage curves, and conducted ICA by differentiating the fitted functions. The SVR method is believed to have good properties: (1) insensitive to measurement noises; (2) robust to data range and size; (3) effective in extracting a signature that shows strong dependence to battery age, making it promising for further online ICA on diagnosing the battery SOH.

However, the SVR method in [35] was not widely propagated for SOH diagnosis after its first publication, because the influences of the key parameters in the SVR algorithm on the IC results have not been clarified. The other researchers may find it difficult to learn the usage of the SVR method, because they may encounter problems such as: (1) how to reduce the computational time of the SVR method; (2) how to judge the accuracy of the curve fitting by the SVR method; (3) does the SVR method fit for lithium-ion cells with other chemistries other than lithium phosphate etc. Those problems hindered a wider application of the SVR method in processing ICA.

This paper dedicates to solve the problems that hindered the wider application of the SVR method in processing ICA. Parametric studies are conducted to investigate the influence of the key parameters in the SVR algorithm on the performance of curve fitting. The relationship between the computational time and the data size is clarified. The influence of the standard deviation σ in the Gaussian kernel on the accuracy of curve fitting has been investigated. The principle for selecting proper σ to avoid over-fitting and under-fitting is proposed. Improvements of the previous SVR algorithm have been proposed by (1) using double σ to adapt to the curve shape with different curvatures; (2) changing the cost function for optimization. The proper parameters are double checked by the degradation data for commercial lithium-ion cells with LFP = $LiFePO_4$ and with NCM = $Li(Ni_xCo_yMn_{1-x-y})O_2$, recommended for further usage of ICA.

2. Experimental

Degradation data are collected to check the validity and universality of the proposed SVR method. The data come from the University of Michigan, Ann Arbor (UoM) and Tsinghua University (THU). The details for the accelerating degradation tests can be found in the marked references. The current listed in the Table 1 refers to that used to generate the voltage curve that needs to be fitted by the SVR method. Figure 1 shows the voltage profiles for the four commercial lithium-ion cells after being cycled by hundreds of times. The voltage profiles in Figure 1 are to be fitted by the SVR method for further derivation of IC curves. Two types of batteries have cathode of LFP, whereas the other two have the cathode with NCM. The mass ratio of Ni:Co:Mn for the batteries with NCM cathode is 1:1:1.

Table 1. Degradation tests of commercial lithium-ion batteries for validating the methodologies. The abbreviation for cathode means: (1) LFP = $LiFePO_4$, (2) NCM = $Li(Ni_xCo_yMn_{1-y})O_2$, (3) LMO = $LiMn_2O_4$; whereas that for anode means: C = carbon or graphite based anode. The voltage ranges comply with the instructions provided by the manufacturers.

Cell	Cathode	Anode	Capacity (Ah)	Voltage (V)	Current	Institute	Refference
A	LFP	C	1.1	3.0–3.55	1/2 C CHA	UoM	[35]
B	NCM	C	48	3.0–4.2	1/3 C CHA	THU	[26]
C	LFP	C	60	2.75–3.6	1/3 C CHA	THU	[26]
D	NCM + LMO	C	20	2.5–4.2	1/2 C DIS	THU	[36]

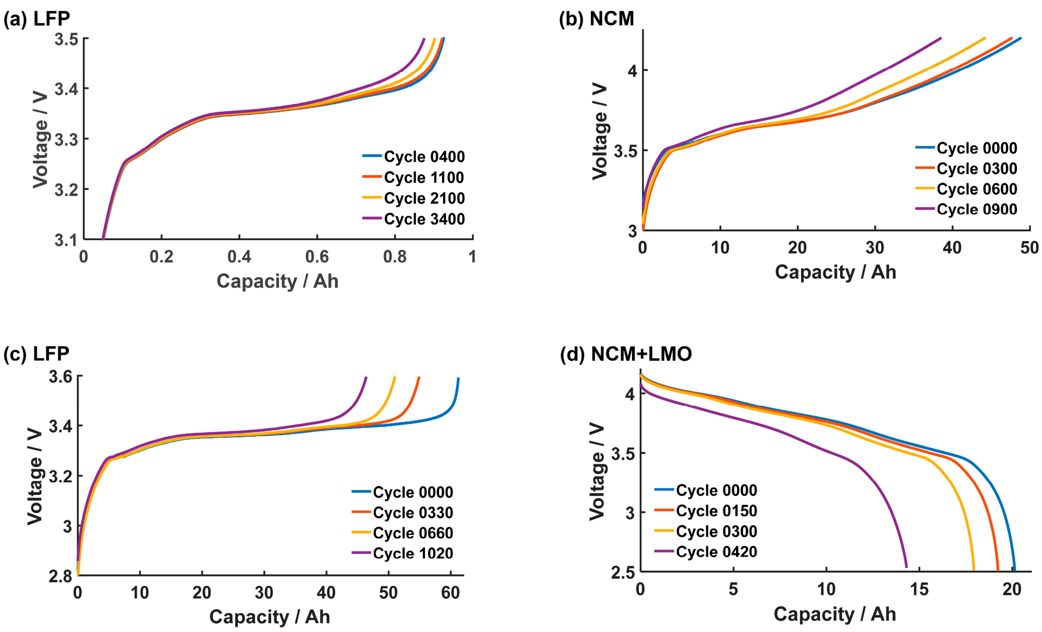

Figure 1. The voltage profile for cells with different cycle numbers. (**a**) Battery A, charge; (**b**) battery B, charge; (**c**) battery C, charge; (**d**) battery D, discharge.

3. Methodology

3.1. The Canonical Form of the Support Vector Regression

The SVR method is an application of the support vector machine in artificial intelligence, which was initially developed by Drucker et al. at the AT&T Bell Laboratories, for data regression [37]. Assume x and y represent the input and output data to be fitted, respectively. Here for battery sampling data, x can be the time τ, the charged/discharged capacity Q, or the state-of-charge SOC, whereas y is the sampled voltage. $X = \{x_1, x_2, x_3 \ldots x_n\}$ and $Y = \{y_1, y_2, y_3 \ldots y_n\}$ are the data set of x and y, respectively. n is the total number of data for curve fitting. For the SVR method, the output data y is fitted using a linearized parametric model:

$$\hat{y} = f(x) = \boldsymbol{\beta}^T \cdot \boldsymbol{\kappa}(x) + \mu = \sum_{k=1}^{n} \beta_k \cdot \kappa(x_k, x) + \mu \tag{1}$$

where $\boldsymbol{\beta} = [\beta_1, \beta_2 \ldots \ \beta_n]^T \in R^{1 \times n}$; μ is the offset constant; $\boldsymbol{\kappa}(x) = [\kappa(x_1, x), \kappa(x_2, x), \ldots \kappa(x_n, x)]^T$ is the vector for kernels. The arbitrary kernel function $\kappa(x_k, x)$ is defined as:

$$\kappa(x_k, x) = \exp\left(\frac{-\|x_k - x\|^2}{2\sigma^2}\right) \tag{2}$$

where σ, which is the standard deviation of the Gaussian kernel, is a preset parameter to control the shape of the $\kappa(x_k, x)$. Define ε as the precision parameter that set some tolerances for the fitting error, and slack variables ξ_k^+ and ξ_k^- to cope with infeasible constraints:

$$\xi_k = \begin{cases} \xi_k^+ = y - \hat{y} - \varepsilon, & (y > \hat{y} + \varepsilon) \\ \xi_k^- = \hat{y} - \varepsilon - y, & (y < \hat{y} - \varepsilon) \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

Figure 2 shows the graphical definition of ε, ξ_k^+ and ξ_k^- for the SVR algorithm. The region of green belt, looking like a tube, indicates that there will be no penalty in the cost function when the fitted curve is quite close to the raw data.

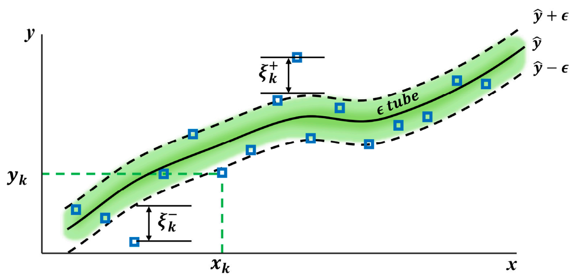

Figure 2. The definition of ε, ξ_k^+ and ξ_k^- for support vector regression.

Then the SVR using l_1 regularization formulates the optimization problem as follows:

$$\begin{aligned} \text{Costfunction\#1}: \quad & \min_{\beta,\mu,\xi^+,\xi^-} \|\boldsymbol{\beta}\|_1 + w \cdot \sum_{k=1}^{n} (\xi_k^- + \xi_k^+) \\ & \text{subjectto} \begin{cases} \hat{y}_k - y_k \le \varepsilon + \xi_k^+ \\ y_k - \hat{y}_k \le \varepsilon + \xi_k^- \\ \xi_k^+ \ge 0 \\ \xi_k^- \ge 0 \end{cases} \end{aligned} \tag{4}$$

where w = 100 is the weighting factor, $||\cdot||_1$ denotes the l_1 norm in the coefficient space.

To further reformulate Equation (4) into a canonical form of Linear Programming (LP) problem, the coefficient β should be decomposed into:

$$\beta_k = \alpha_k^+ - \alpha_k^-, \ |\beta_k| = \alpha_k^+ + \alpha_k^- \tag{5}$$

where α_k^+ and α_k^- are nonnegative and satisfy $\alpha_k^+ \cdot \alpha_k^- = 0$.

Then the SVR problem in Equation (4) can be further reformulated into the carnonical form of LP problem:

$$\begin{aligned} & \min_{c} \boldsymbol{c}^T \cdot \boldsymbol{z} \\ & \text{subject to}: \ A \cdot \boldsymbol{z} \le \boldsymbol{b} \end{aligned} \tag{6}$$

where

$$\begin{cases} \boldsymbol{c} = \left(\underbrace{1,1,\ldots 1}_{2n},\underbrace{w,w,\ldots w}_{2n},0\right)^T \\ A = \begin{pmatrix} K & -K & -I & 0 & 1 \\ -K & K & 0 & -I & -1 \end{pmatrix} \\ \boldsymbol{z} = \begin{pmatrix} \boldsymbol{\alpha}^+ \\ \boldsymbol{\alpha}^- \\ \boldsymbol{\xi}^+ \\ \boldsymbol{\xi}^- \\ \mu \end{pmatrix}, (\boldsymbol{\alpha}^+,\boldsymbol{\alpha}^-,\boldsymbol{\xi}^+,\boldsymbol{\xi}^- \geq 0) \\ \boldsymbol{b} = \begin{pmatrix} \varepsilon + \boldsymbol{y} \\ \varepsilon - \boldsymbol{y} \end{pmatrix} \end{cases} \quad (7)$$

and

$$\begin{cases} K = \begin{bmatrix} \kappa(x_1,x_1) & \kappa(x_1,x_2) & \ldots & \kappa(x_1,x_n) \\ \kappa(x_2,x_1) & \kappa(x_2,x_2) & \ldots & \kappa(x_2,x_n) \\ \ldots & \ldots & \ldots & \ldots \\ \kappa(x_n,x_1) & \kappa(x_n,x_2) & \ldots & \kappa(x_n,x_n) \end{bmatrix} \\ \boldsymbol{y} = (y_1,y_2,\ldots y_n)^T \\ \boldsymbol{\alpha}^+ = \left(\alpha_1^+,\alpha_2^+,\ldots \alpha_n^+\right)^T \\ \boldsymbol{\alpha}^- = \left(\alpha_1^-,\alpha_2^-,\ldots \alpha_n^-\right)^T \\ \boldsymbol{\xi}^+ = \left(\xi_1^+,\xi_2^+,\ldots \xi_n^+\right)^T \\ \boldsymbol{\xi}^- = \left(\xi_1^-,\xi_2^-,\ldots \xi_n^-\right)^T \end{cases} \quad (8)$$

The optimization problem in Equation (6) can be solved using the function *linprog* in MATLAB (2016a®). The computational complexity of the problem is $\Theta(n^2)$. n is usually large for battery test data, therefore it should be resampled with a distance of D, to downsize the optimization to n/D, and the computational complexity will be reduced to $\Theta(n^2/D^2)$. For example, assume that the raw data R is $R = \{r_1, r_2, \ldots, r_n\}$, then the resampled data will be $S = \{s_j \mid s_j = r_{j \cdot D}, j = 1,2, \ldots, [n/D]\} = \{r_D, r_{2\times D}, \ldots, r_{[n/D]\times D}\}$.

The optimal result usually give near-zero value for most of the β_i, and those β_i that are much larger than zero (>10^{-4}) are regarded as significant, and corresponding x_i is called the support vector sv_i, of which the total number is N_{sv}. Therefore the model for curve fitting is built as:

$$\hat{y} = f(x) = \boldsymbol{\beta}^T \cdot \boldsymbol{\kappa}(x) + \mu = \sum_{k=1}^{N_{sv}} \beta_k \cdot \kappa(x_k, x) + \mu \quad (9)$$

Figure 3 illustrates the selection of sv_i according to the optimal solution found by the Cost Function #1 in Equation (4). There are 22 sv_i that has $\beta_i > 10^{-4}$, therefore $N_{sv} = 22$ is not that large comparing with the original data size $n \sim 10^4$. Note that some sv_i locate close with each other, therefore it looks that there are only 15 separate sv_i (marked with cyan diamonds) as in Figure 3.

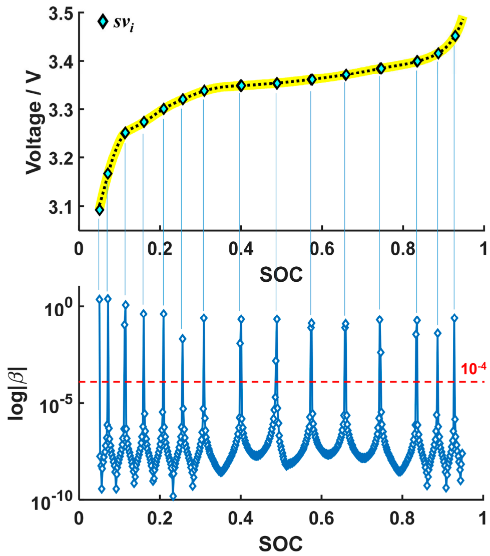

Figure 3. The selection of support vector with significant coefficient β, for Battery A, σ = 0.06, D = 100.

3.2. Double σ to Enhance the Quality of Curve Fitting

In Reference [35] a fixed σ is used for the SVR method, however, sometimes the SVR method cannot always make good fit for a full range of *SOC* with a fixed σ in application. The reason for the problem is that the shape of the kernel $\kappa(x_k, x)$ is mainly controlled by σ (as shown in Figure 4a), although its coefficient β can regulate the shape a little bit. If the shape of the original voltage matches that of the kernel, the SVR method with a single fixed σ will have a good fit result, otherwise in some *SOC*s the SVR method will return overfitted or underfitted results. On this account, the SVR method in Reference [35] has been upgraded for fitting the data with different curvatures. That is the SVR algorithm using double σ in the Gaussian kernel.

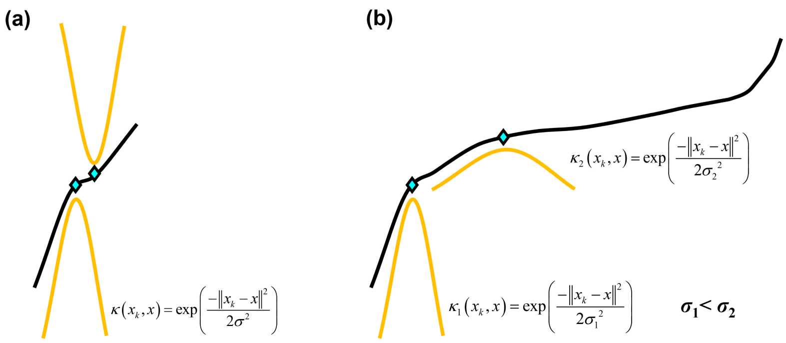

Figure 4. Using double σ to enhance the quality of curve fitting, (**a**) SVR algorithm with single σ; (**b**) SVR algorithm with double σ.

Reformulating the Equation (1) with two different kernels $\boldsymbol{\kappa}_1(x)$ and $\boldsymbol{\kappa}_2(x)$:

$$\hat{y} = f(x) = \boldsymbol{\beta}_1{}^T \cdot \boldsymbol{\kappa}_1(x) + \boldsymbol{\beta}_2{}^T \cdot \boldsymbol{\kappa}_2(x) + \mu = \sum_{k=1}^{n} \beta_{1,k} \cdot \kappa_1(x_k, x) + \sum_{k=1}^{n} \beta_{2,k} \cdot \kappa_2(x_k, x) + \mu \tag{10}$$

where the Gaussian kernels yields to:

$$\begin{cases} \kappa_1(x_k, x) = \exp\left(\frac{-\|x_k - x\|^2}{2{\sigma_1}^2}\right) \\ \kappa_2(x_k, x) = \exp\left(\frac{-\|x_k - x\|^2}{2{\sigma_2}^2}\right) \end{cases} \tag{11}$$

Here there are two different σ, for which $\sigma_1 < \sigma_2$. The kernel with a smaller σ is used to fit the voltage segment with big curvatures, whereas that with a larger σ is used to fit those with small curvatures, as shown in Figure 4b. The correlated functions for optimal solution should be changed accordingly, however, they will not be listed here to save contents.

3.3. Criterion for Evaluating the Quality of Curve Fitting by SVR

The accuracy of the SVR method should be judged in order to fulfill further improvement. Here we are proposing a new criterion using linear approximation of discretized sampling data to judge the accuracy of the curve fitting. Further discussion on evaluating the quality of the curve fitting by SVR will be based on the criterion of linear approximation. Figure 5 illustrates the criterion for evaluating the quality of curve fitting by SVR, according to the property of the discretized sampling data with stairs and noises. Figure 5a shows a common case of the voltage sampling data, with clear stair-like shape caused by digital sampling. The actual differentiation $y' = \mathrm{d}y/\mathrm{d}x$ at Y_i is approximately the slope of the red line, or tanθ, as calculated in Equation (12):

$$y\prime = \frac{\mathrm{d}y}{\mathrm{d}x} = \tan\theta \approx \frac{1}{N_{Y_i}} \cdot \frac{\delta Y}{\delta X} \tag{12}$$

where N_{Yi} is the number of points that satisfy $y_k = Y_i$ in the output data sequence of $Y = \{y_1, y_2, y_3 \ldots y_n\}$, δY and δX are the resolution of sampling for the output and input data, respectively. Here δY = 1 mV and δX = 0.1 s, representing a sampled data sequence in a real case. Moreover, for Figure 5a, the y' at $y_k = Y_i$ should be 10^{-3}, taking that N_{Yi} = 10. Therefore in an ideal case, the SVR method should return a result of $\hat{y}\prime = f\prime(x)$ close to 10^{-3} for all the x_k that have $y_k = Y_i$. Figure 5b demonstrates the ideal curve fitting case for the SVR method. Both the $\hat{y} = f(x)$ and $\hat{y}\prime = f\prime(x)$ (yellow line in Figure 5b) looks smooth when they cross those points with $y_k = Y_i$, and the fitted differentiation $\hat{y}\prime = f\prime(x)$ is close to the slope tanθ that we have calculated from Figure 5a.

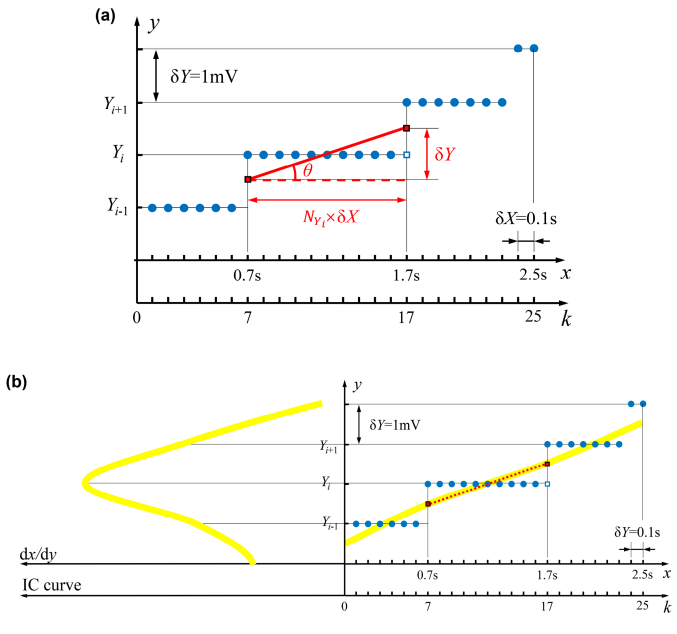

Figure 5. *Cont.*

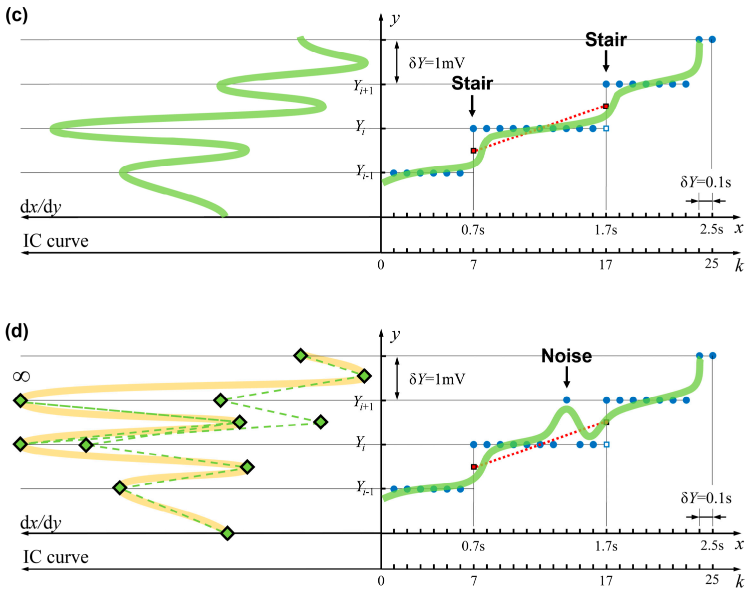

Figure 5. Establishment of criterion for evaluating the quality of curve fitting by SVR. (**a**) The approximate linear differentiation that the SVR method dedicates to capture. (**b**) A case of good fit for the SVR method. (**c**) A case of bad fit, caused by overfitting the shape of stairs for the discretized data. (**d**) A case of bad fit, caused by overfitting the shape of noise.

Figure 5c illustrates an overfitting case caused by the stair-like sampled voltage. Peaks can be observed in the dx/dy curves for each level of voltage stairs, if the SVR method tracks them closely. Figure 5d illustrates an overfitting case caused by noises in the sampled data. Extra peaks that superimposed on the peaks in Figure 5c commence in the dx/dy curves, if the SVR method tries to follow the noise. If the SVR method keeps tracking the shape of the sampled voltage with Cost Function #1 (Equation (4)), the over-fitting cases in Figure 5c,d will be inevitable. Therefore we try to propose a new cost function to regulate the curve-fitting of the SVR method.

The selection of δY is important for the shape of the IC curve calculated by the linear approximation. Define δR as the real resolution of the discrete sampling, δY should be the multiples of δR ($\delta Y = K{\cdot}\delta R$, K is an integer) in order to avoid fluctuations caused by unfair counting. Unfair counting means that if $\delta Y \neq K{\cdot}\delta R$, sometimes δY contains $K{\cdot}\delta R$, while sometimes δY contains $(K + 1){\cdot}\delta R$, leading to inevitable fluctuations in the processed results for linear approximation. As the sampling accuracies of voltage in this paper are all 1 mV, δY should be the multiples of 1 mV for calculating the linear approximation. Larger δY can reduce the fluctuations in the differentiations results by the linear approximation, as shown in Figure 6. But too large δY will lead to loss of useful information of the original data. The tradeoff should be considered when selecting a proper δY. Therefore in this paper, δY = 2 mV for the cell with LFP cathode, whereas δY = 5 mV for the cell with NCM cathode.

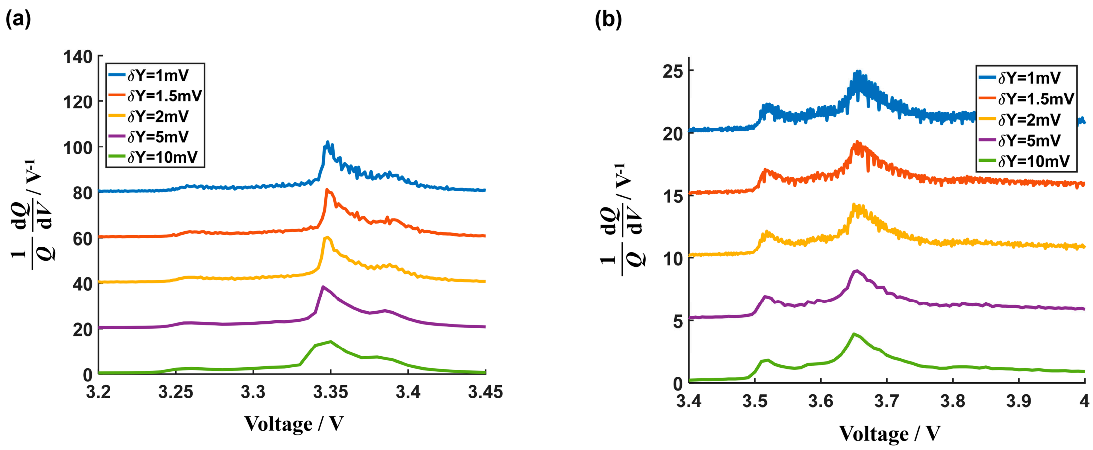

Figure 6. The selection of proper sampling resolution for the approximate linear differentiation in Figure 5a. (**a**) Cell A with LFP cathode. (**b**) Cell B with NCM cathode.

3.4. Changing the Cost Functions to Improve the Accuracy of Incremental Capacity Analysis

The slack variables that are critical in the Cost Function #1 (Equation (4)) come from Equation (3), with the sampled sequence y as the data to be fitted. y', which is linearly approximated in Equation (12), can be used as a new reference for defining the new slack variables ς_k in Equation (13).

$$\varsigma_k = \begin{cases} \varsigma_k^+ = y'^* - \hat{y}\prime - \gamma, & (y'^* > \hat{y}\prime + \gamma) \\ \varsigma_k^- = \hat{y}\prime - \gamma - y'^*, & (y'^* < \hat{y}\prime - \gamma) \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

where the reference differentiation y' is marked as y'^*, in order to be distinguished from the real differentiation of y. In Equation (13), γ is the precision parameter that sets tolerance for the fitting error, and slack variables ς_k^+ and ς_k^- are used to cope with infeasible constraints. Furthermore, the fitting function should be:

$$\hat{y}\prime = \boldsymbol{\beta}^T \cdot \boldsymbol{\kappa}\prime(x) = \sum_{k=1}^{n} \beta_k \cdot \kappa\prime(x_k, x) \tag{14}$$

where μ is omitted due to the first-order differentiation, κ' is the first-order differentiation of the Gaussian kernel functions, as defined in Equation (15).

$$\kappa\prime(x_k, x) = -\frac{x - x_k}{\sigma^2} \cdot \kappa(x_k, x) \tag{15}$$

Then the Cost Function #2, which is established according to the deviations in the first-order differentiation can be built by Equation (16).

$$\begin{aligned} \text{CostFunction\#2}: \quad & \min_{\beta,\varsigma^+,\varsigma^-} \|\boldsymbol{\beta}\|_1 + \lambda \cdot \sum_{k=1}^{n} (\varsigma_k^- + \varsigma_k^+) \\ & \text{subject to} \begin{cases} \hat{y\prime}_k - y\prime_k^* \leq \gamma + \varsigma_k^+ \\ y\prime_k^* - \hat{y\prime}_k \leq \gamma + \varsigma_k^- \\ \varsigma_k^+ \geq 0 \\ \varsigma_k^- \geq 0 \end{cases} \end{aligned} \tag{16}$$

where $\lambda = 100$ is the weighting factor. Similarly with that in Equation (5), the coefficient β is decomposed into nonnegative α_k^+ and α_k^-, satisfying $\alpha_k^+ \cdot \alpha_k^- = 0$, in order to form a carnonical form of LP problem. Moreover, the SVR problem with Cost Function #2 can be reformulated into the carnonical form of the LP problem as in Equation (6). And the vectors and matrices are changed from Equation (7) into Equation (17):

$$\begin{cases} c = \left(\underbrace{1, 1, \ldots 1}_{2n}, \underbrace{\lambda, \lambda, \ldots \lambda}_{2n} \right)^T \\ A = \begin{pmatrix} K\prime & -K\prime & -I & 0 \\ -K\prime & K\prime & 0 & -I \end{pmatrix} \\ z = \begin{pmatrix} \boldsymbol{\alpha}^+ \\ \boldsymbol{\alpha}^- \\ \boldsymbol{\varsigma}^+ \\ \boldsymbol{\varsigma}^- \end{pmatrix}, (\boldsymbol{\alpha}^+, \boldsymbol{\alpha}^-, \boldsymbol{\varsigma}^+, \boldsymbol{\varsigma}^- \geq 0) \\ \boldsymbol{b} = \begin{pmatrix} \gamma + \boldsymbol{y}\prime \\ \gamma - \boldsymbol{y}\prime \end{pmatrix} \end{cases} \tag{17}$$

and

$$\begin{cases} K\prime = \begin{bmatrix} \kappa\prime(x_1, x_1) & \kappa\prime(x_1, x_2) & \ldots & \kappa\prime(x_1, x_n) \\ \kappa\prime(x_2, x_1) & \kappa\prime(x_2, x_2) & \ldots & \kappa\prime(x_2, x_n) \\ \ldots & \ldots & \ldots & \ldots \\ \kappa\prime(x_n, x_1) & \kappa\prime(x_n, x_2) & \ldots & \kappa\prime(x_n, x_n) \end{bmatrix} \\ \boldsymbol{y}' = (y\prime_1, y\prime_2, \ldots y\prime_n)^T \\ \boldsymbol{\alpha}^+ = (\alpha_1^+, \alpha_2^+, \ldots \alpha_n^+)^T \\ \boldsymbol{\alpha}^- = (\alpha_1^-, \alpha_2^-, \ldots \alpha_n^-)^T \\ \boldsymbol{\varsigma}^+ = (\varsigma_1^+, \varsigma_2^+, \ldots \varsigma_n^+)^T \\ \boldsymbol{\varsigma}^- = (\varsigma_1^-, \varsigma_2^-, \ldots \varsigma_n^-)^T \end{cases} \tag{18}$$

Solving the LP problem in Equations (16)–(18), the fitted first-order differentiation can be:

$$\hat{y}\prime = \boldsymbol{\beta}^T \cdot \boldsymbol{\kappa}(x) = \sum_{k=1}^{N_{sv}} \beta_k \cdot \kappa(x_k, x) \tag{19}$$

which is the direct fitting results of the IC by the SVR algorithm.

4. Result and Discussions

The performances of curve-fitting by the SVR algorithm will be discussed in this section. The reasons of over-fitting and under-fitting have been studied in details, and correlated improvements to avoid those problems are discussed.

4.1. The Influence of the Data Length on the Performances of the SVR Algorithm

Parametric studies are conducted on the influence of the data length n on the goodness of fitting and on the computational time of the SVR algorithm. Here we choose $\sigma = 0.20$ for the SVR algorithm, identical with that set in Reference. [35]. Figure 7a shows that as D increases, the data length n/D used as the input of the SVR algorithm decreases from approximately 3000 to 100. There is always under-fitting at the *SOC* near 0% and 100% for all data lengths. And the degree of under-fitting will become worse for the algorithm with shorter data lengths. The phenomenon is similar for both the Cell A with LFP cathode and Cell B with NCM cathode, because similar phenomenon can be seen in both Figure 7a,b. Figure 7c shows that the IC curves look similar for $200 < n/D < 3000$ for Cell A with LFP cathode. However, for $n/D < 200$, the characteristic IC peak located between 3.35 V and 3.4 V disappears. Therefore downsampling with too few points of data is not acceptable to keep the accuracy of IC curves. Similar phenomenon can be observed in Figure 7d for Cell B with NCM cathode.

The data length also determines the computational time of the SVR algorithm, as shown in Figure 7e. As the computational complexity for the LP problem in Equation (6) is $\Theta(n^2)$, the downsampling will help save much time in computation by reducing the computational complexity to $\Theta\ (n^2/D^2)$. The computational times for the SVR algorithm that is used to conduct curve-fitting for cell with LFP and NCM cathode have been compared in Figure 7e. The workstation that is used for the SVR algorithm has an Intel® Core™ i7-6820HK CPU @2.70 GHz with a RAM of 16.0 GB, and the version of the MATLAB® used for computation is R2016a. The computational times display shapes of polynomial n^2 as in Figure 7e as expected, therefore n must be well-controlled within a lower range to save the time used for curve fitting.

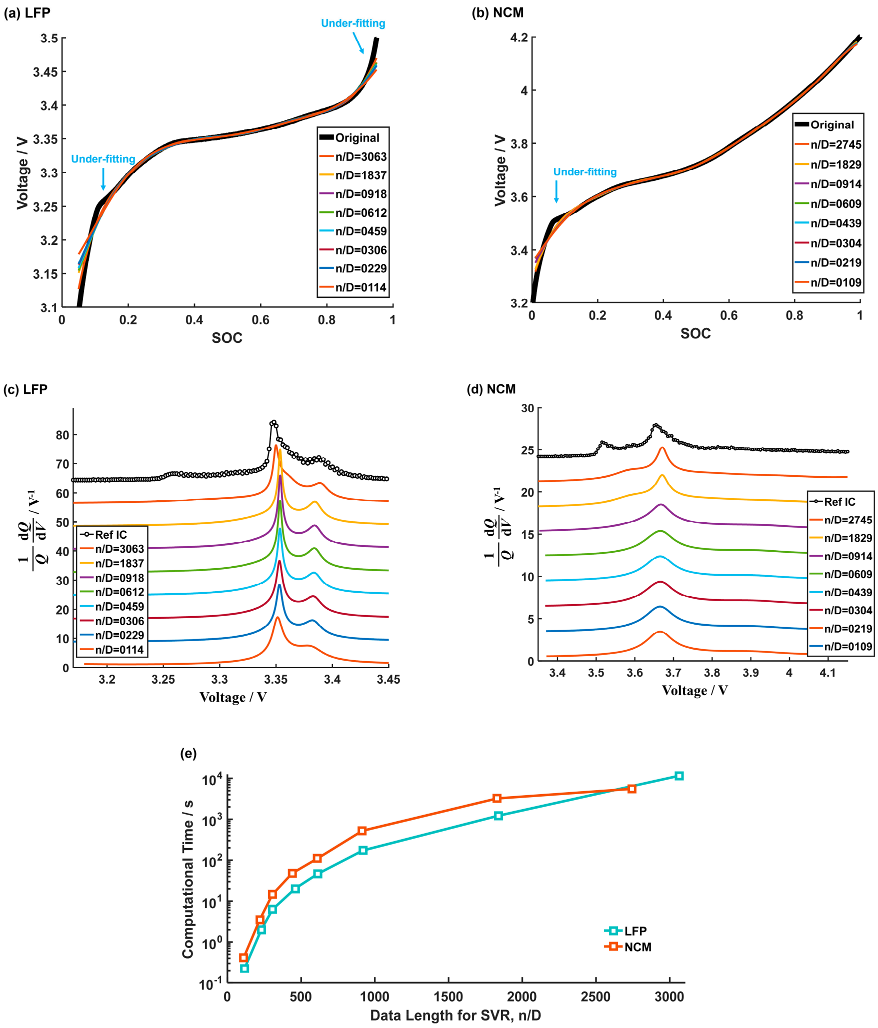

Figure 7. The relationship between the data length and the performances of the SVR algorithm. (**a**) The fitting of the voltage for Cell A with LFP cathode. (**b**) The fitting of the voltage for Cell B with NCM cathode. (**c**) The fitting results of the IC curves for Cell A with LFP cathode. (**d**) The fitting results of the IC curves for Cell B with NCM cathode. (**e**) Computational time versus the data length for the SVR algorithm.

In summary, there is a trade-off between the computational time and the accuracy of curve fitting when selecting a proper data length n for the SVR algorithm. Nevertheless, the accuracy of the SVR algorithm in fitting the IC curve is not majorly determined by n as long as there is enough data that can reflect the overall shape of the voltage curve. Later discussions show that the accuracy of the SVR algorithm is much more influenced by the σ and the cost function. Improvements will be fulfilled by selecting proper σ and cost function. Therefore the computational time has higher priority when selecting the divider D for downsampling. The data length n/D is controlled within 400–500 points for further discussion, balancing both the computational time and accuracy of curve fitting.

4.2. The Influence of the σ in the Gaussian Kernel on the Performance of the SVR Algorithm

The σ in the Gaussian kernel (Equation (2)) controls the shape of individual support vectors, and the fitted curve $\hat{y}$ is a combination of several kernels, therefore σ influences the accuracy of curve fitting by the SVR algorithm. Figure 8a,b illustrate the curve fitting results of the voltage for Cell A with LFP cathode and Cell B with NCM cathode, respectively. The problem of under-fitting as shown in Figure 7a,b have been solved by selecting smaller σs. For Cell A with LFP cathode, σ = 0.06 is much better than σ = 0.20, whereas for Cell B with NCM cathode, σ = 0.08 is much better than σ = 0.20. The reason that why the SVR algorithm with smaller σ performs better than that with large σ is that the basic curvatures formed by support vectors with smaller σ can be finer than those with larger σ. The number of support vectors (N_{sv}) will increase for the SVR algorithm with smaller σ, because the algorithm is trying to capture the finer details of the voltage curves.

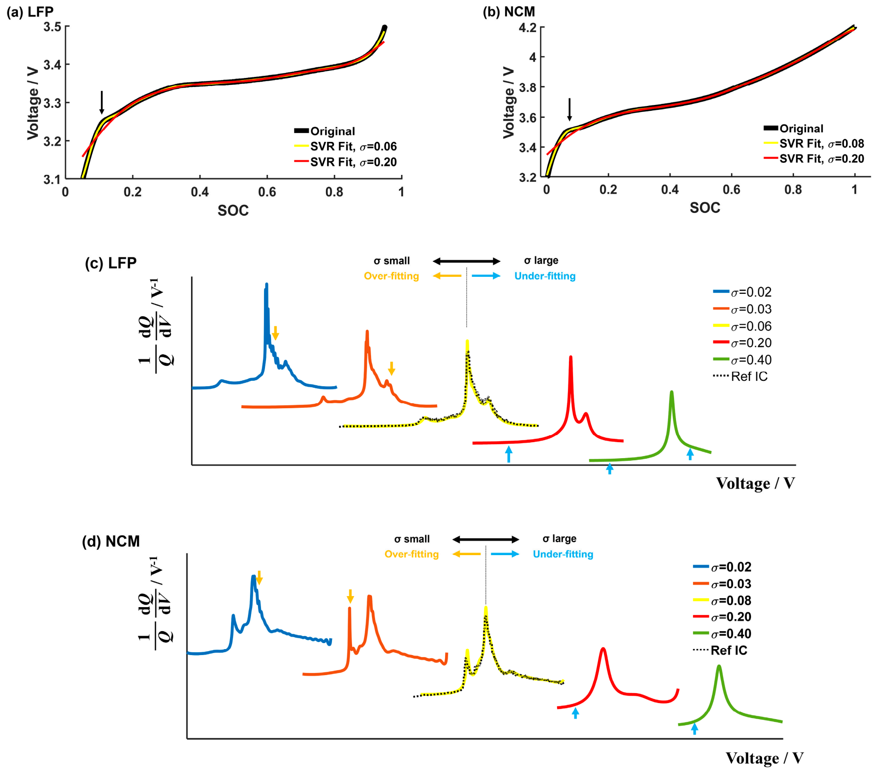

Figure 8. The influence of the σ in the Gaussian kernel on the accuracy of curve fitting. (**a**) The fitting results of the voltage curve for Cell A with LFP cathode, σ = 0.06 and 0.20. (**b**) The fitting results of the voltage curve for Cell B with NCM cathode, σ = 0.08 and 0.20. (**c**) The fitting results of the IC curve for Cell A with LFP cathode, from small σ = 0.02 to large σ = 0.40. (**d**) The fitting results of the IC curve for Cell B with NCM cathode, from small σ = 0.02 to large σ = 0.40.

However, overfitting will appear if too small σ is chosen in the SVR algorithm both for the Cell A with LFP cathode and for the Cell B with NCM cathode, as shown in Figure 8c,d. The mechanism of overfitting in the IC curves can be explained by Figure 5c,d. The SVR algorithm has seemingly stronger capability to capture finer details of the voltage curves with smaller σ, therefore they will reflect the stairs and noises in the IC results. Overfitting is inevitable when extremely small σ is set in the SVR algorithm. Here when $\sigma < 0.04$, overfitting will occur for both cells with LFP cathode and

NCM cathode. On the other hand, too large σ will lead to under-fitting in the voltage curves especially for SOC near 0% and 100%, as we already know from Figure 8a,b. Some IC peaks will disappear for the SVR algorithm with $\sigma > 0.10$. Therefore if we wish to use single σ in the SVR algorithm, a proper range can be $0.04 < \sigma < 0.10$. Here we choose $\sigma = 0.06$ for Cell A with LFP cathode, and $\sigma = 0.08$ for Cell B with NCM cathode in the following discussions.

4.3. Double σ to Enhance the Accuracy of the SVR Algorithm

Overfitting or underfitting is still possible to occur because the shape of the voltage curves will change during degradation after cycling. If overfitting or under-fitting only appears when curve fitting is conducted for the cycled data, it means that the fixed σ for the fresh cell no longer fits for the changed curvature in the cycled voltage curve. Large σ fits for voltage curve with small curvatures, whereas small σ fits for that with big curvatures. As the σ is preset and fixed in Section 4.2, the problem caused by cell degradation is usually possible for the SVR algorithm. One way to solve the problem is to use double σ, of which the smaller one is used to fit the arc shape with big curvature, whereas the larger one to fit that with small curvature. The mathematical formulation of the SVR algorithm using double σ can be referred in Section 3.2.

Figure 9a,b show the curve fitting results for Cell A and Cell B using the SVR algorithm with double σ, respectively. The support vectors with larger σ is prone to be located at the SOC near 0% and 100%, also at the position with small curvature. The sitting of the support vectors with larger σ ($\sigma_2 = 0.20$) help smooth out the curve fitting results, with lower possibility to have overfitting problems, as those shown in Figure 8. Moreover, the voltage curve is still mainly "supported" by the smaller σ ($\sigma_1 = 0.06$ for LFP cell, and 0.08 for NCM cell) with high accuracy. Figure 9c,d display the IC curves derived from fitting voltage for Cell A with LFP cathode and Cell B with NCM cathode, respectively. The overfitting problem still occurs at minor positions for the SVR algorithm with single σ, as in the magnified figures in Figure 9c,d. Note that without magnification, the overfitting problem in the IC curves cannot be noticed from Figure 8c,d. The fluctuations in the IC curves have been successfully smoothed out with the help of double σ in the SVR algorithm. Hence, utilizing double σ, one smaller for the accuracy of fitting and another bigger one for avoiding overfitting, can substantially improve the quality of voltage curve fitting for lithium-ion batteries. Up to now the SVR algorithm in Reference [35] has been herein improved with better accuracy in curve fitting and with lower possibility of overfitting. The SVR algorithm in Reference [35] cannot capture all the voltage curvatures at all SOCs, but with double σ, the problem has been solved. After checking the curve fitting results for the cells before and after degradation, we recommend that the double σ can be set as $\{\sigma_1 = 0.06, \sigma_2 = 0.20\}$ for LFP cells, and $\{\sigma_1 = 0.08, \sigma_2 = 0.20\}$ for NCM cells.

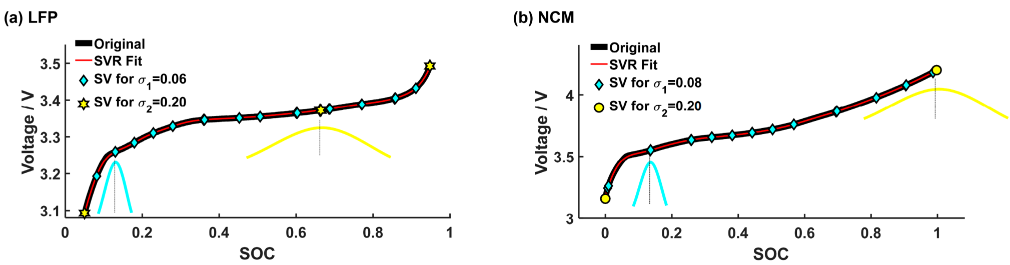

Figure 9. *Cont.*

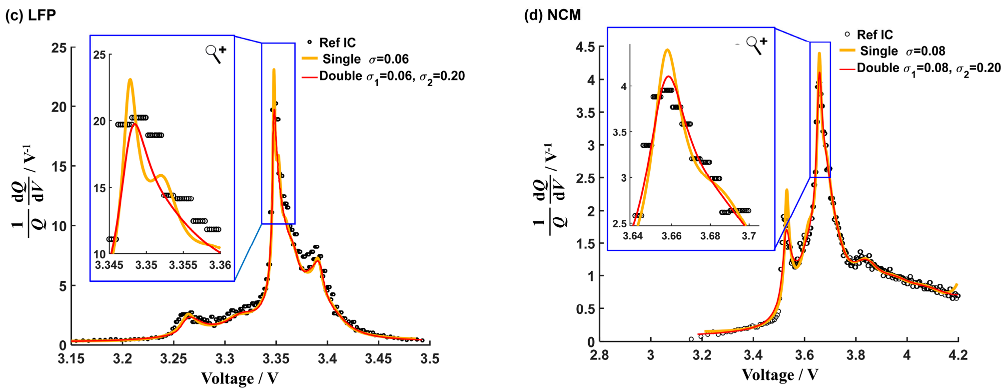

Figure 9. Using double σ to improve the accuracy of the SVR algorithm. (**a**) The curve fitting of the voltage for Cell A with LFP cathode. (**b**) The curve fitting of the voltage for Cell B with NCM cathode. (**c**) The fitting result of the IC curve for Cell A with LFP cathode. (**d**) The fitting result of the IC curve for Cell B with NCM cathode.

4.4. Changing the Cost Function to Improve the Accuracy of ICA Using the SVR Algorithm

Up to now, the accuracy of curve fitting by the SVR algorithm has been greatly improved comparing with that in Reference [35]. The following deficiencies of the SVR algorithm have been overcome: (1) the deficiency in accurately fitting the voltage throughout the full SOC ranges (from 0% to 100%); (2) the deficiency of possible overfitting or under-fitting that can be displayed in the IC curves; (3) the questionable capability of the SVR algorithm to fit the voltage curve for lithium-ion cells with NCM cathode. In this section, we are trying to further seek finer fitting of the voltage curves of lithium-ion batteries using the SVR algorithm. And a criterion for judging the accuracy of the curve fitting in the view of the differentiation has been proposed as in Equation (20):

$$RMSE = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left(y\prime_i^* - \hat{y}\prime_i\right)^2} \quad (20)$$

where RMSE means the overall root-mean-square-error for the differentiation $\hat{y}\prime$ calculated from the fitted function $\hat{y}$, the subscript i is the index of the data, and n is the data length.

The Cost Function #2 (Equation [15], which directly regulates the LP algorithm by the voltage differentiation rather than the voltage, is utilized to further improve the quality of the SVR algorithm as discussed in Sections 4.2 and 4.2. Not only the influence of the changes in the cost functions, but also the performances of choosing single σ and double σ with different cost functions have been compared. Therefore the performances of four kinds of SVR algorithms have been compared, as listed in Table 2. The discussion of the performances focus on the (1) fitting accuracy, with or without overfitting/under-fitting; (2) whether the algorithm can return a continuous function of the fitted curves, $y = f(x)$ for the voltage curve and $y' = f'(x)$ for the IC curve. Recommendations follow the correlated discussions for further usage of the algorithm in ICA.

Table 2. Comparison of the SVR algorithms with different parametric settings.

Algorithm		Fitting Accuracy	$y = f(x)$	$y' = f'(x)$	Further Aging ICA
σ	**Cost Function**				
Single σ	#1	Fair	Explicit	Explicit	Not Recommended
Single σ	#2	Better	No	Explicit	Recommended
Double σ	#1	Good	Explicit	Explicit	Recommended
Double σ	#2	Better	No	Explicit	Recommended

Figure 10a,b compares the IC curves returned by the four different SVR algorithms for Cell A with LFP cathode and Cell B with NCM cathode, respectively. The reference IC curves (the black dotted line marked as Ref IC in the Figure 10a,b) come from y' approximated by Equation (12). Obvious deviations can be seen for the SVR algorithm with single σ and Cost Function #1 for both cases, called as "Fair" in Table 2. Improvements can be seen for the SVR algorithm with double σ and Cost Function #1, as we expected from the discussion in Section 4.3, called as "Good" in Table 2. Further improvements can be seen for the SVR algorithm with Cost Function #2, regardless of the number of σ that is used in the SVR algorithm, called as "Better" in Table 2. The further improvements in the IC curve using Cost Function #2 come from that the SVR algorithm directly regulates the estimation error by the Ref IC, which we believe to be the best linear approximation by Equation (12). However, the price paid for using Cost Function #2 is that we cannot rebuild the voltage curve $y = f(x)$, because the offset constant μ has been erased from the function of $y' = f'(x)$ when calculating the first-order differentiation.

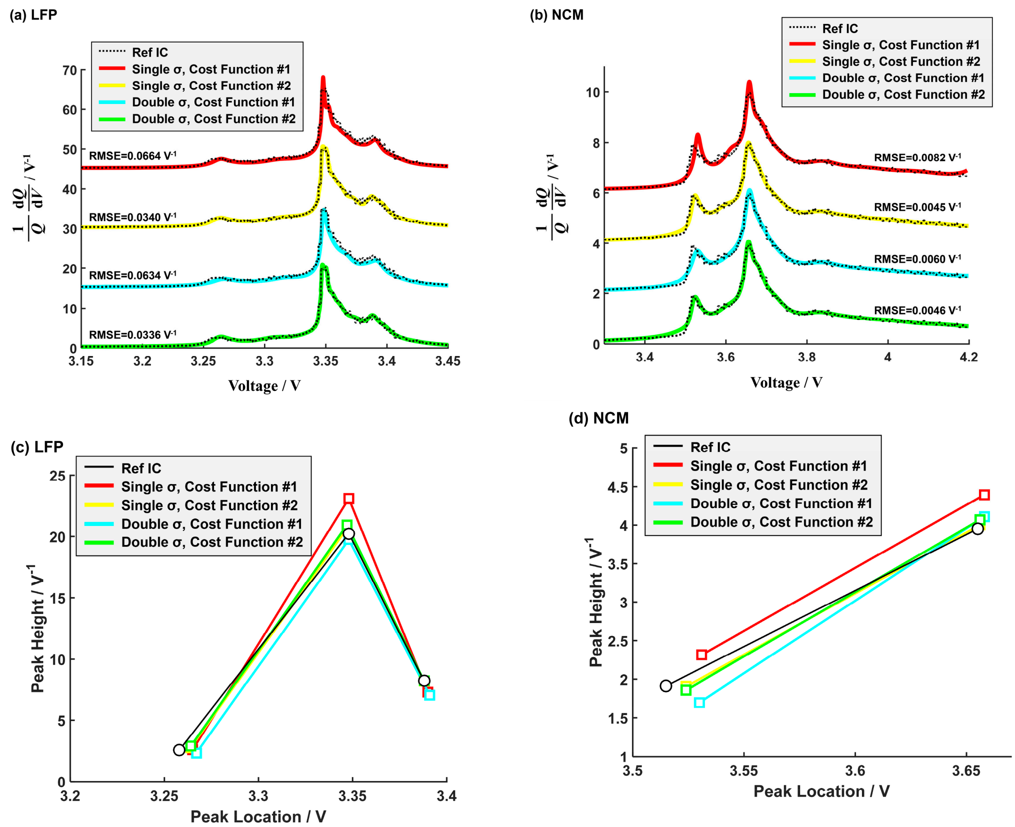

Figure 10. Comparisons of the fitting results of IC curves for the SVR algorithms with different settings. Comparison of the fitted IC curves for (**a**) Cell A with LFP cathode; (**b**) Cell B with NCM cathode. Comparison of the locations and heights of the peaks in the IC curves for (**c**) Cell A with LFP cathode; (**d**) Cell B with NCM cathode.

The SVR algorithms with "single σ + Cost Function #2" and "double σ + Cost Function #1" are recommended for further aging analysis based on ICA. The SVR algorithm with "single σ + Cost Function #1" is not recommended because its fair accuracy in fitting the IC curve. The SVR algorithm with "double σ + Cost Function #2" is also not recommended, because it has higher computational complexity than that with "single σ + Cost Function #2", while the two algorithms have similar accuracy in the view of *RMSE*. Moreover, if we focus on the location and height of the characteristic peaks in the IC curves as shown in Figure 10c,d, the SVR algorithm with "single σ + Cost Function #2" and with "double σ + Cost Function #2" also show similar results. Although the accuracy of the SVR algorithm with "double σ + Cost Function #1" is not as good as that with "single σ + Cost Function

#2", it is still recommended because it can provide explicit continuous functions of the voltage curves, which is important to conduct further research on the curvatures of the voltage at any arbitrary *SOC*.

4.5. ICA for Characterizing the Degradation of Lithium-ion Batteries Using the SVR Algorithm

The SVR algorithm with "single σ + Cost Function #2", that with "double σ + Cost Function #1", and that with "double σ + Cost Function #2" are further used to perform ICA to characterize the degradation of commercial lithium-ion batteries, as shown in Figure 11. Table 3 collects the statistics of features of the characteristic peaks. The accuracy of the IC results calculated by the SVR algorithm with "single σ + Cost Function #2" and "double σ + Cost Function #2" are quite similar, therefore we only show the results for "single σ + Cost Function #2" to save contents.

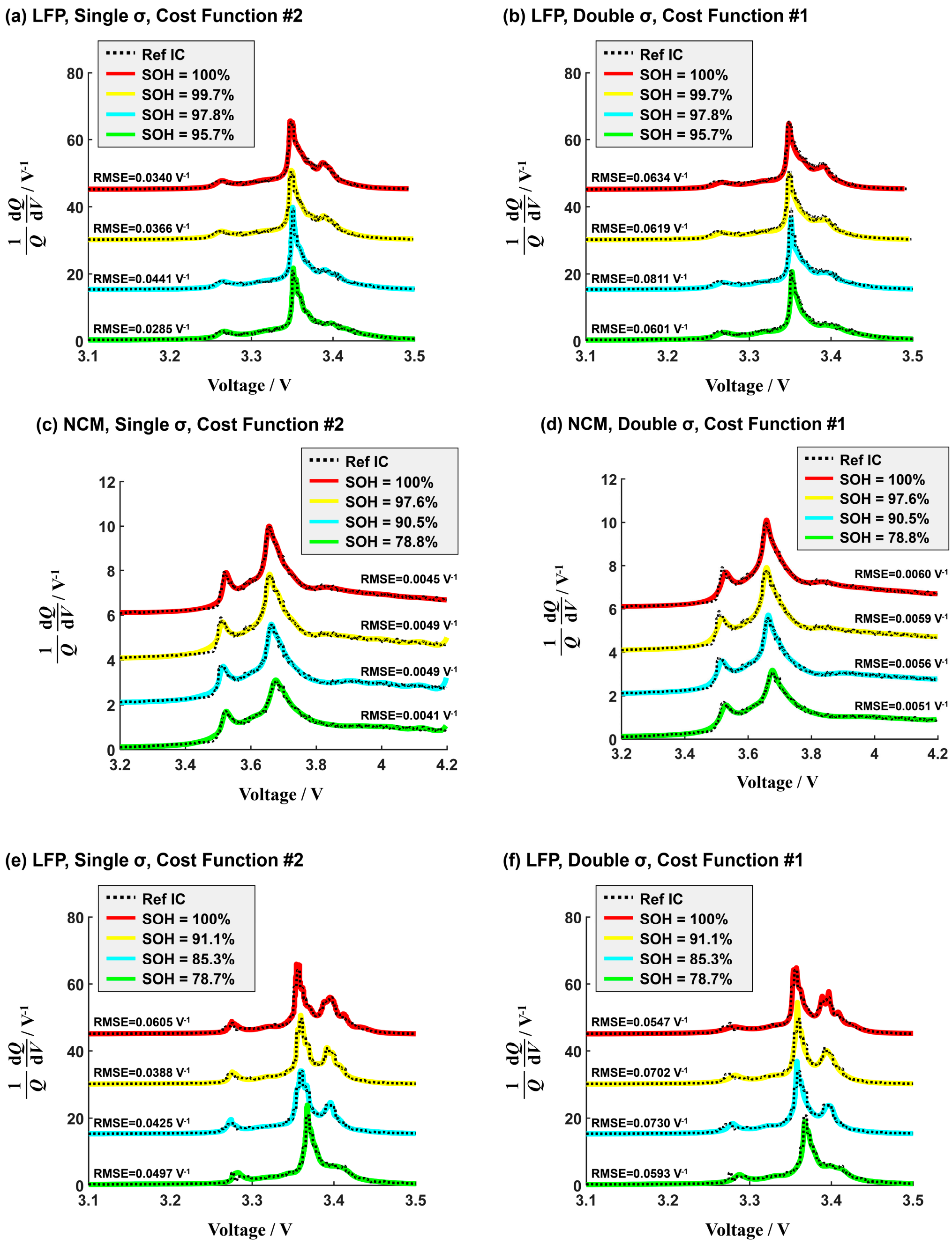

Figure 11. *Cont.*

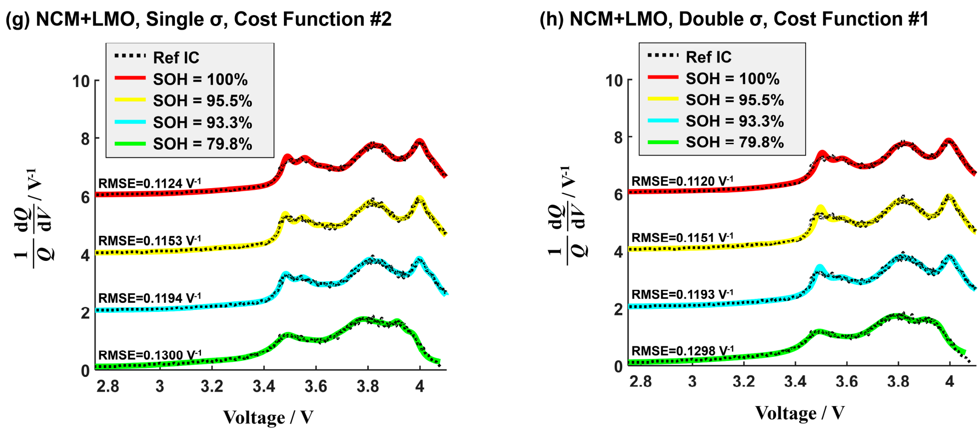

Figure 11. ICA for characterizing the degradation of lithium-ion batteries using the SVR algorithm. (**a**) Cell A with LFP cathode, single σ + Cost Function #2; (**b**) Cell A with LFP cathode, double σ + Cost Function #1; (**c**) Cell B with LFP cathode, single σ + Cost Function #2; (**d**) Cell B with NCM cathode, double σ + Cost Function #1; (**e**) Cell C with LFP cathode, single σ + Cost Function #2; (**f**) Cell C with LFP cathode, double σ + Cost Function #1; (**g**) Cell D with NCM+LMO cathode, single σ + Cost Function #2; (**h**) Cell D with NCM+LMO cathode, double σ + Cost Function #1.

Figure 11 shows that the IC curves calculated by the two SVR algorithms are both close to the Ref IC calculated by the linear approximation in Equation (12). The *RMSE* for the SVR algorithm with "single σ + Cost Function #2" is usually slightly smaller than that with "double σ + Cost Function #1", indicating that Cost Function #2 is dominating the fitting accuracy of the IC curves. Using fixed σ in the SVR algorithm to fit the data throughout the full life cycle is acceptable, because the *RMSE* changes a little at different SOHs, for all the cells in Figure 11. Moreover, $\{\sigma_1 = 0.06, \sigma_2 = 0.20\}$ is good for cells with LFP cathode, because the fitting results for Cell A and Cell C are quite similar. Similarly, $\{\sigma_1 = 0.08, \sigma_2 = 0.20\}$ is good for the cells with NCM cathode, because the fitting results for Cell B and Cell D are quite similar. Other combinations of $\{\sigma_1, \sigma_2\}$ are also possible, readers who have interest could check according to the discussions made in Sections 4.2 and 4.2.

However, problem emerges when we try to found quantitative relationship between the SOH and the IC peaks, especially for the SVR algorithm with Cost Function #1. Quantified features of IC peaks as shown in Figure 12, including the peak location, the peak height, and the integrated area surrounding the peak have been used for establishing relationships between the IC peaks and the SOHs of lithium-ion cells in order to build online SOH estimation algorithms. In Reference [35], quantified relationship has been built between the peak height and the battery SOH. Table 3 shows the accuracy of the peak locations and peak heights using the SVR algorithm with different settings. The calculation of the relative error is conducted from the data in Figure 11, with the reference from the linear approximation by Equation (12). The characteristic peak is picked as long as the peak height increases/decreases monotonically with the SOH decreasing. Although the accuracy of the peak location can be precise with error less than 0.5%, the error in the peak height of the SVR algorithm with "double σ + Cost Function #1" can be as large as 15%–16%, whereas that with "single σ + Cost Function #2" can be up to 8%. The accuracy of the SVR algorithm with "double σ + Cost Function #1" is not acceptable for further quantified ICA analysis, if the peak height is regarded as an indicator.

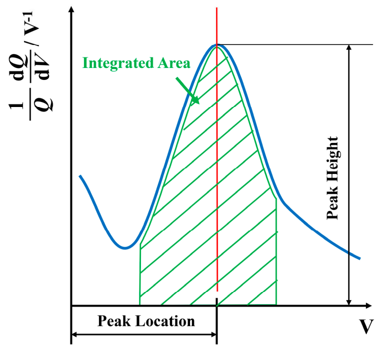

Figure 12. Quantifying the features of IC peak for further online SOH estimation.

Table 3. The accuracy of the recommended SVR algorithms in capturing the feature of the characteristic peaks (peak location and height) in the IC curves.

Cell	Cathode	SOH	Characteristic Peak							
			Reference IC (Linear Approximation)		Relative Error					
					SVR Single σ Cost Function #2		SVR Double σ Cost Function #1		SVR Double σ Cost Function #2	
			Location /V	Height /V^{-1}	Location	Height	Location	Height	Location	Height
A	LFP	100%	3.388	8.205	0%	−0.23%	+0.09%	−14.45%	0%	+0.04%
		99.7%	3.388	7.344	+0.03%	+0.07%	+0.12%	−16.60%	+0.03%	−0.07%
		97.8%	3.388	6.328	+0.06%	+0.03%	+0.09%	−14.84%	+0.06%	−0.30%
		95.7%	3.392	5.573	+0.18%	−4.25%	0%	−9.54%	+0.15%	−4.19%
B	NCM	100%	3.655	3.953	+0.03%	+1.24%	+0.08%	+3.92%	+0.03%	+2.93%
		97.6%	3.655	3.731	+0.05%	+2.68%	+0.08%	+4.90%	0%	+3.03%
		90.5%	3.660	3.563	+0.08%	+0.93%	+0.11%	+4.15%	0%	+0.28%
		78.8%	3.675	3.117	+0.03%	+0%	+0.05%	+2.44%	0%	+0.19%
C	LFP	100%	3.394	11.02	+0.03%	+0.73%	+0.09%	+15.70%	+0.09%	+6.62%
		91.1%	3.390	10.77	+0.06%	+3.25%	+0.15%	−6.78%	+0.24%	−0.74%
		85.3%	3.390	8.873	+0.18%	+8.58%	+0.21%	−1.78%	+0.09%	+5.42%
		78.7%	3.398	6.344	−0.15%	+5.74%	0%	−6.45%	−0.09%	−6.76%
D	NCM+LMO	100%	3.990	1.875	+0.23%	+1.39%	+0.13%	−0.32%	+0.28%	−2.29%
		95.5%	3.990	1.931	+0.20%	+0.73%	+0.13%	−1.40%	+0.25%	−1.50%
		93.3%	3.990	1.823	+0.18%	+1.43%	+0.25%	−1.26%	+0.25%	+0.66%
		79.8%	3.916	1.670	−0.08%	+0.72%	−0.38%	−2.51%	+0.20%	−0.30%

Nevertheless, as the height of the peak calculated by the linear approximation might be influenced by the noise in the raw data, sometimes we can alter to the integrated area surrounding the peak, as the shaded area marked in Figure 12. The errors of the fitted IC curves can be well controlled under 3% through the full cycle life for the SVR algorithm with "single σ + Cost Function #2" for all kinds of commercial lithium-ion batteries, after introducing the integrated area near the IC peak as an SOH indicator. Table 4 lists that most of the error is less than 1%, unless the cell reaches the end of its life. However, the SVR algorithm with "double σ + Cost Function #1" still has an error as large as 10% for Cell A with LFP cells, although it has less error (up to 5%) than using the peak height as the SOH indicator.

In summary, the SVR algorithm with Cost Function #2 performs much better than those with Cost Function #1. And the SVR algorithm with Cost Function #2 is promising to be further used for online SOH estimation, helping build quantified relationship between the IC peaks and the battery SOH. For the SVR algorithm with "single σ + Cost Function #2", σ = 0.06 is good for cells with LFP cathode, whereas σ = 0.08 is good for cells with NCM cathode. A smaller σ can also cause overfitting in the results of the SVR algorithm with "single σ + Cost Function #2", the reason of which is similar with that discussed in Section 4.2. The SVR algorithm with Cost Function #1 also has its own advantage that it can rebuild the voltage curve, for which that with Cost Function #2 cannot. The SVR algorithm with "double σ + Cost Function #1" can perform well in many cases, although sometimes the error can be as large as 10% for fitting the peak heights. One should choose proper settings of the SVR algorithm for specific applications, and the discussions in this paper can guide the selections of proper parameters for the SVR algorithms.

Table 4. The accuracy of the recommended SVR algorithms in capturing the features of the characteristic peaks (height integration near the peak) in the IC curves.

Cell	Cathode	SOH	Integrated Area Near the Peak	SVR Single σ + Cost Function #2 Relative Error	SVR Double σ + Cost Function #1 Relative Error
A	LFP	100%	[3.380, 3.400] V	+0.61%	−9.96%
		99.7%	[3.380, 3.400] V	−0.76%	−9.89%
		97.8%	[3.380, 3.400] V	−0.90%	−8.65%
		95.7%	[3.380, 3.400] V	+1.85%	−10.65%
B	NCM	100%	[3.620, 3.750] V	−0.34%	−0.15%
		97.6%	[3.620, 3.750] V	−0.52%	−0.14%
		90.5%	[3.620, 3.750] V	0.39%	+0.10%
		78.8%	[3.620, 3.750] V	0.68%	+0.21%
C	LFP	100%	[3.380, 3.410] V	+0.40%	+0.78%
		91.1%	[3.380, 3.410] V	−0.33%	−2.62%
		85.3%	[3.380, 3.410] V	−1.99%	−3.45%
		78.7%	[3.380, 3.410] V	−2.76%	−5.29%
D	NCM+LMO	100%	[3.910, 4.050] V	+0.49%	+0.04%
		95.5%	[3.910, 4.050] V	+0.25%	+0.19%
		93.3%	[3.910, 4.050] V	−0.14%	−0.21%
		79.8%	[3.880, 4.020] V	−0.82%	−0.31%

5. Conclusions

This paper studies the performances of the SVR algorithm on the ICA for commercial lithium-ion batteries. The SVR algorithm, which serves a curve fitting method using Gaussian kernel, can smooth the discrete sampled voltage data with continuous functions, thereby making further differentiations applicable for the ICA. Parametric studies have been conducted to improve the performance of the SVR algorithm in fitting the stair like sampled voltage from four kinds of commercial lithium-ion batteries. The computational time can be reduced by downsampling the raw data with proper resampling distance D. The possible problem of overfitting and under-fitting can be solved by choosing proper standard deviation σ in the Gaussian kernel in the SVR algorithm. Smaller σ is prone to cause overfitting, whereas larger σ is likely to bring under-fitting, therefore an SVR algorithm with double σ (one large and one small) is proposed to compensate the disadvantages of using single σ. The cost function, which is previously set to regulate the error in voltage fitting, has been further modified to directly regulate the error in the differentiations. The new cost function greatly improves the accuracy of the SVR algorithm in calculating the IC curves, making it promising to be further applied in online SOH estimation. The maximum error in the characteristic IC peaks calculated by the SVR algorithm with the new cost function can be less than 3% for all the data during the full cycle life of all the four kinds of lithium-ion batteries. After viewing all the curve fitting results, $\sigma = 0.06$ is good for cells with LFP cathode, whereas $\sigma = 0.08$ is good for cells with NCM cathode.

Author Contributions: Conceptualization, X.H. (Xiangming He) and M.O.; Data curation, C.W., D.R. and X.H. (Xuebing Han); Investigation, X.F.; Methodology, X.F. and C.W.; Project administration, X.H. (Xiangming He) and M.O.; Resources, L.L.; Supervision, X.H. (Xuebing Han) and M.O.; Validation, L.W. and L.L.; Writing-original draft, X.F.; Writing-review & editing, L.W.

Funding: This work is supported by the [National Key R&D Program of China] grant number [2018YFB0104404]; [National Natural Science Foundation of China] grant number [51706117] and [U1564205]; by [China Postdoctoral Science Foundation] grant number [2017M610086]. The first author thanks the support of "Young Elite Scientist Sponsorship Program" from China Association for Science and Technology.

Acknowledgments: Xuning Feng appreciates the beneficial discussion with Gregory J. Offer with Imperial College London, and Shriram Santhanagopalan with National Renewable Energy Laboratory.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hoog, J.D.; Jaguemont, J.; Abdel-Monem, M.; Bossche, P.V.D.; Mierlo, J.V.; Omar, N. Combining an electrothermal and impedance aging model to investigate thermal degradation caused by fast charging. *Energies* **2018**, *11*, 804. [CrossRef]
2. Lai, X.; Zheng, Y.; Sun, T. A comparative study of different equivalent circuit models for estimating state-of-charge of lithium-ion batteries. *Electrochimica Acta* **2018**, *259*, 566–577. [CrossRef]
3. Hu, X.; Martinez, C.M.; Yang, Y. Charging, power management, and battery degradation mitigation in plug-ion hybrid electric vehicles: A unified cost-optimal approach. *Mech. Syst. Signal Process.* **2017**, *87*, 4–16. [CrossRef]
4. Kong, X.; Zheng, Y.; Ouyang, M.; Lu, L.; Li, J.; Zhang, Z. Fault diagnosis and quantitative analysis of micro-short circuits for lithium-ion batteries in battery packs. *J. Power Sources* **2018**, *395*, 358–368. [CrossRef]
5. Zou, C.; Hu, X.; Wei, Z.; Wik, T.; Egardt, B. Electrochemical estimation and control for lithium-ion battery health-aware fast charging. *IEEE Trans. Ind. Electron.* **2018**, *65*, 6635–6645. [CrossRef]
6. Huang, S.C.; Tseng, K.H.; Liang, J.W.; Chang, C.L.; Pecht, M.G. An online SOC and SOH model for lithium-ion batteries. *Energies* **2017**, *10*, 512. [CrossRef]
7. Zheng, Y.; Ouyang, M.; Han, X.; Lu, L.; Li, J. Investigating the error sources of the online state of charge estimation methods for lithium-ion batteries in electric vehicles. *J. Power Sources* **2018**, *377*, 161–188. [CrossRef]
8. Li, X.; Jiang, J.; Wang, L.Y.; Chen, D.; Zhang, Y.; Zhang, C. A capacity model based on charging process for state of health estimation of lithium ion batteries. *Appl. Energy* **2016**, *177*, 537–543. [CrossRef]
9. Ouyang, M.; Feng, X.; Han, X.; Lu, L.; Li, Z.; He, X. A dynamic capacity degradation model and its applications considering varying load for a large format Li-ion battery. *Appl. Energy* **2016**, *165*, 48–59. [CrossRef]
10. Yu, J.; Mo, B.; Tang, D.; Yang, J.; Wan, J.; Liu, J. Indirect state-of-health estimation for lithium-ion batteries under randomized use. *Energies* **2017**, *10*, 2012. [CrossRef]
11. Hu, X.; Jiang, J.; Cao, D.; Egardt, B. Battery health prognosis for electric vehicles using sample entropy and sparse Bayesian predictive modeling. *IEEE Trans. Ind. Electron.* **2016**, *63*, 2645–2654. [CrossRef]
12. Mathew, M.; Janhunen, S.; Rashid, M.; Long, F.; Fowler, M. Comparative analysis of lithium-ion battery resistance estimation techniques for battery management systems. *Energies* **2018**, *11*, 1490. [CrossRef]
13. Bao, Y.; Dong, W.; Wang, D. Online internal resistance measurement application in lithium ion battery capacity and state of charge estimation. *Energies* **2018**, *11*, 1073. [CrossRef]
14. Love, C.T.; Dubarry, M.; Reshetenko, T.; Devie, A.; Spinner, N.; Swider-Lyons, K.E.; Rocheleau, R. Lithium-ion cell fault detection by single-point impedance diagnostic and degradation mechanism validation for series-wired batteries cycled at 0 °C. *Energies* **2018**, *11*, 834. [CrossRef]
15. Zou, Y.; Hu, X.; Ma, H.; Li, S.E. Combined state of charge and state of health estimation over lithium-ion battery cell cycle lifespan for electric vehicles. *J. Power Sources* **2015**, *273*, 793–803. [CrossRef]
16. Shen, P.; Ouyang, M.; Lu, L.; Li, J.; Feng, X. The co-estimation of state of charge, state of health and state of function for lithium-ion batteries in electric vehicles. *IEEE Trans. Veh. Technol.* **2018**, *67*, 92–103. [CrossRef]
17. Wu, T.H.; Moo, C.S. State-of-charge estimation with state-of-health calibration for lithium-ion batteries. *Energies* **2017**, *10*, 987. [CrossRef]
18. Balewski, L.; Brenet, J.P. A new method for the study of the electrochemical reactivity of manganese dioxide. *Electrochem. Technol.* **1967**, *5*, 527–531.
19. Dubarry, M.; Svoboda, V.; Hwu, R.; Liaw, B.Y. Incremental capacity analysis and close-to-equilibrium OCV measurements to quantify capacity fade in commercial rechargeable lithium batteries. *Electrochem. Solid-State Lett.* **2006**, *9*, A454–A457. [CrossRef]
20. Dubarry, M.; Liaw, B.Y. Identify capacity fading mechanism in a commercial $LiFePO_4$ cell. *Electrochem. Technol.* **2009**, *194*, 541–549. [CrossRef]

21. Zhang, C.; Yan, F.; Du, C.; Kang, J.; Turkson, R.F. Evaluating the degradation mechanism and state of health of LiFePO4 lithium-ion batteries in real-world plug-ion hybrid electric vehicles application for different ageing paths. *Energies* **2017**, *10*, 110. [CrossRef]
22. Han, X.; Ouyang, M.; Lu, L.; Li, J.; Zheng, Y.; Li, Z. A comparative study of commercial lithium ion battery cycle life in electrical vehicle: Aging mechanism identification. *J. Power Sources* **2014**, *251*, 38–54. [CrossRef]
23. Dubarry, M.; Svoboda, V.; Hwu, R.; Liaw, B.Y. Capacity and power fading mechanism identification formo a commercial cell evaluation. *J. Power Sources* **2007**, *165*, 566–572. [CrossRef]
24. Gao, Y.; Jiang, J.; Zhang, C.; Zhang, W.; Ma, Z.; Jiang, Y. Lithium-ion battery aging mechanisms and life model under different charging stresses. *J. Power Sources* **2017**, *356*, 103–114. [CrossRef]
25. Jiang, Y.; Jiang, J.; Zhang, C.; Zhang, W.; Gao, Y.; Guo, Q. Recognition of battery aging variations for $LiFePO_4$ batteries in 2nd use applications combining incremental capacity analysis and statistical approaches. *J. Power Sources* **2017**, *360*, 180–188. [CrossRef]
26. Feng, X.; Li, J.; Ouyang, M.; Lu, L.; Li, J.; He, X. Using probability density function to evaluate the state of health of lithium-ion batteries. *J. Power Sources* **2013**, *232*, 209–218. [CrossRef]
27. Weng, C.; Feng, X.; Sun, J.; Peng, H. State-of-health monitoring of lithium-ion battery modules and packs via incremental capacity peak tracking. *Appl. Energy* **2016**, *180*, 360–388. [CrossRef]
28. Christopheren, J.P.; Shaw, S.R. Using radial basis functions to approximate battery differential capacity and differential voltage. *J. Power Sources* **2010**, *195*, 1225–1234. [CrossRef]
29. Samad, N.A.; Kim, Y.; Siegel, J.B.; Stefanopoulou, A.G. Battery capacity fading estimation using a force-based incremental capacity analysis. *J. Electrochem. Soc.* **2016**, *163*, A1584–A1594. [CrossRef]
30. Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squared procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [CrossRef]
31. He, Y.; Shen, J.; Shen, J.; Ma, Z. Embedding monotonicity in the construction of polynomial open-circuit voltage model for lithium-ion batteries: A semi-infinite programming formulation approach. *Ind. Eng. Chem. Res.* **2015**, *54*, 3167–3174. [CrossRef]
32. Hu, X.; Li, S.E.; Yang, Y. Advanced machine learning approach for lithium-ion battery state estimation in electric vehicles. *IEEE Trans. Transport. Electrif.* **2016**, *2*, 140–149. [CrossRef]
33. Zhang, C.; Jiang, J.; Zhang, W.; Wang, Y.; Sharkh, S.M.; Xiong, R. A novel data-driven fast capacity estimation of spent electric vehicle lithium-ion batteries. *Energies* **2014**, *7*, 8076–8094. [CrossRef]
34. Peng, Y.; Hou, Y.; Song, Y.; Pang, J.; Liu, D. Lithium-ion battery prognostics with hybrid Gaussian process function regression. *Energies* **2018**, *11*, 1420. [CrossRef]
35. Weng, C.; Cui, Y.; Sun, J.; Peng, H. On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression. *J. Power Sources* **2013**, *235*, 36–44. [CrossRef]
36. Ren, D.; Lu, L.; Ouyang, M.; Feng, X.; Li, J.; Han, X. Degradation identification of individual components in the $Li_yNi_{1/3}Co_{1/3}Mn_{1/3}O_2$-$Li_yMn_2O_4$ blended cathode for large format lithium ion battery. *Energy Procedia* **2017**, *105*, 2698–2704. [CrossRef]
37. Drucker, H.; Burges, C.J.C.; Kaufman, L.; Smola, A.; Vapnik, V. Support vector regression machines. In Proceedings of the 9th International Conference on Neural Information Processing Systems, Denver, CO, USA, 3–5 December 1996; pp. 155–161.