State of charge classification for lithium-ion batteries using impedance based features

Marian Patrik Felder and Jürgen Götze

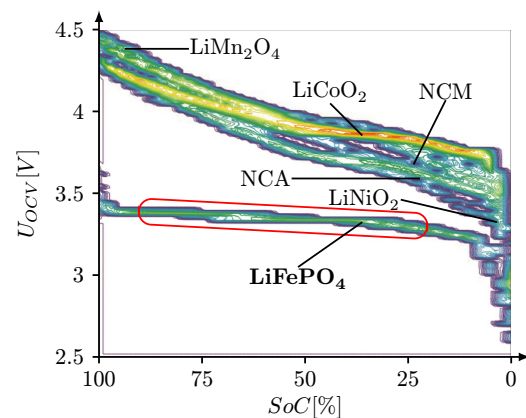Information Processing Lab, TU Dortmund University, 44221 Dortmund, Germany

Correspondence to: Marian Patrik Felder (marian.felder@tu-dortmund.de) and Jürgen Götze
(juergen.goetze@tu-dortmund.de)

Abstract. Currently, the electrification of the drive train of passenger cars takes place, and the task of obtaining precise knowledge about the condition of the on board batteries gains importance. Due to a flat open circuit voltage (OCV) to state of charge (SoC) characteristic of lithium ion batteries, methods employed in applications with other cell chemistries cannot be adapted. Exploiting the higher significance of the impedance for state estimation for that chemistry, new impedance based features are proposed by this work. To evaluate the suitability of these features, simulations have been conducted using a simplified on-board power supply net as excitation source. The simulation outcome has been investigated regarding the cross correlation factor r_{xy} and in a polynomial regression scenario. The results of the simulations show a best case error below 1 % SoC, which is 3 percentage points lower than using terminal voltage and impedance. When increasing the measurement uncertainty, the difference remains around 2 percent points.

Figure 1. OCV curves over SoC for different cell chemistries (Keller, 2013).

1 Introduction

A main challenge to the current generation of engineers is the successful transition from fossil fuels to the use of regenerative energy sources for stationary automotive applications. In Germany, the electrification of drives is propelled by CO_2 reduction laws (European Parliament, 2014), demanding an average emission of $95\,g\,CO_2\,km^{-1}$ by all new cars by 2021, as well as the objective of one million electric vehicles (EVs) by 2020 and six million EVs by 2030 (Second Report by The German National Platform for Electric Mobility, 2011). Due to electrification of the drive train in the car, the importance of batteries as power storage increases. Lithium ion batter-

ies are promising prospects to achieve the set goals, because of the high energy density compared to other commercially available cells.

Condition monitoring of the EVs battery is an important task in the automotive field: The state of charge (SoC) is the "fuel gauge" to the battery. Precise knowledge of the SoC does not only provide information about the remaining distance coverable, but also enables the system to keep a safety buffer for the function of security systems in case of need.

For conventional cell chemistries, the open circuit voltage (OCV) was a considerable measure for the SoC (Richardson et al., 2012; Feng et al., 2015; Chang, 2013, Sect. 3.1.1 and sources therein). This does not apply to Li ion batteries, as the OCV characteristic is very flat, with a slope only a few $1\,mV$ for each decade of % SoC (see Fig. 1, with $LiFePO_4$ as extreme example). Furthermore, the OCV curve based meth-

ods require a rest period, which length is difficult to quantify and challenging to implement in xEV applications as clocks and some other loads cannot be turned off entirely (Piller et al., 2001). As an alternative, the impedance has shown to be a significant SoC-sensitive feature (Piller et al., 2001; Klotz et al., 2011; Schmidt et al., 2011).

The Taylor Fourier transform (TFT) is a method originally used in stationary power line applications, where it is used to identify faults from abnormal mains frequency properties (de la O Serna and Rodríguez, 2007). As a novelty in this work, the TFT is used to extract impedance based features. Furthermore, the aptness of the features regarding battery state monitoring is investigated in an exemplary use case. Especially for low noise, the results are promising.

The rest of this work is structured as follows: Sect. 2 will describe the underlying theory of the signal model (Sect. 2.1) and the extraction algorithm (Sect. 2.2). In Sect. 3, the simulation setup is presented. The 4th section covers the investigations about the suitability of the new features, explicitly by correlation coefficient analysis (Sect. 4.1) and an exemplary use case in a polynomial regression (Sect. 4.2). Finally, Sect. 5 concludes this work with a summary and suggestions regarding future work.

2 Theory

In this section, the signal model used in the investigations is explained in detail. After that, the Taylor Fourier transform is described as it is used for feature extraction. The signal model and derivation of the TFT follow de la O Serna and Rodríguez (2007).

2.1 Signal model

Let $s(n)$ be a real valued, sampled oscillating signal with sample index n, signal frequency f, sampling rate f_s, amplitude $a(n)$ and phase offset $\varphi(n)$.

$$s(n) = a(n)\cos\left(2\pi\frac{f}{f_s}n + \varphi(n)\right) \tag{1}$$

In Eq. (2), $a(n)$ and $\varphi(n)$ can be gathered in a complex pointer called dynamic phasor $p(n)$, with z^* denoting the complex conjugate of $z \in \mathbb{C}$. As an abbreviation, $\omega = 2\pi\frac{f}{f_s}$ is used for the normalized angular frequency corresponding to f and f_s. This yields an equation, where the amplitude and phase shift are gathered in the variable $p(t)$.

$$s(n) = \frac{1}{2}\left(a(n)e^{j(2\pi\frac{f}{f_s}n+\varphi(n))} + a(n)e^{-j(2\pi\frac{f}{f_s}n+\varphi(n))}\right) \tag{2}$$

$$= \frac{1}{2}\left(p(n)e^{j(\omega n)} + p^*(n)e^{-j(\omega n)}\right) \tag{3}$$

In case of the battery, measurements of the voltage will not oscillate with zero mean, in contrast to the assumptions by de la O Serna and Rodríguez (2007). In this case, the measured signal $x(n)$ with bias $c \in \mathbb{R}$ would be

$$x(n) = s(n) + c \tag{4}$$

For brevity and simplicity, $s(n)$ has no bias c. This can be achieved by subtracting the mean value c from the measured data.
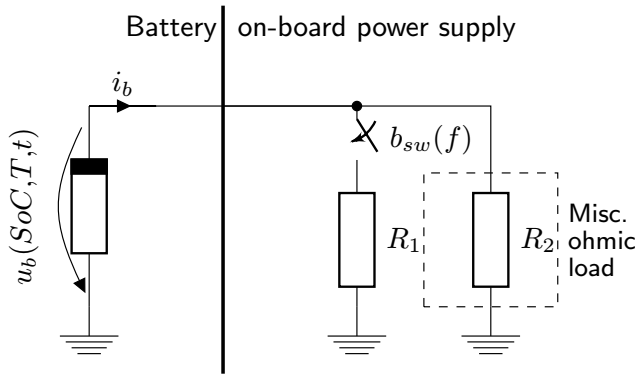
2.2 Taylor Fourier transform

The Taylor Fourier transform can be used to extract $p(n)$ and its derivatives. To follow the evolution of these parameters, the time signal is filtered using a sliding window. Let $s_k(n)$ be the kth window of the sampled signal $s(n)$ from above with odd window length N and $N_h = \lfloor N \rfloor$ its rounded down half. Then, $\bar{s}_k^{(2)}(n)$ is an approximation for $s_k(n)$ with respect to a truncated Taylor series of p_k of second order, $p_k^{(2)}$ (de la O Serna and Rodríguez, 2007).

$$\bar{s}_k(n)\Re\left\{\left[p_{0,k}^{(2)} + p_{1,k}^{(2)}(n - N_h) + p_{2,k}^{(2)}(n - N_h)^2\right]e^{j\omega n}\right\} \tag{5}$$

To create a system of linear equations, rewrite Eq. (5) using the complex conjugate like in Eq. (3). Let the time samples be vector $\boldsymbol{n}$, the exponential functions vector $\boldsymbol{b}$, then matrix $\mathbf{D}_k(\boldsymbol{n}, \boldsymbol{b})$ is a function of both. Furthermore, vector $\boldsymbol{p}_k$ consists of factors $p_{i,k}$. This altogether yields Eq. (6).

With the pseudo inverse $\mathbf{D}_k^{\#} = \left(\mathbf{D}_k^T\mathbf{D}_k\right)^{-1}\mathbf{D}_k^T$ of $\mathbf{D}_k$, this over-determined system of equations can be solved in least squares sense. $\mathbf{D}_k^T$ is the transpose of $\mathbf{D}_k$. This finally yields an approximation of the parameters $\bar{p}_k$.

$$\underbrace{\begin{bmatrix} \bar{s}_k^{(2)}(0) \\ \vdots \\ \bar{s}_k^{(2)}(N_h) \\ \vdots \\ \bar{s}_k^{(2)}(n) \\ \vdots \\ \bar{s}_k^{(2)}(N-1) \end{bmatrix}}_{} = \frac{1}{2}\underbrace{\begin{bmatrix} N_h^2 e^{j\omega N_h} & -N_h e^{j\omega N_h} & e^{j\omega N_h} & e^{-j\omega N_h} & -N_h e^{-j\omega N_h} & N_h^2 e^{-j\omega N_h} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ n^2 e^{-j\omega n} & n e^{-j\omega n} & e^{-j\omega n} & e^{j\omega n} & n e^{j\omega n} & n^2 e^{j\omega n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ N_h^2 e^{-j\omega N_h} & N_h e^{-j\omega N_h} & e^{-j\omega N_h} & e^{j\omega N_h} & N_h e^{j\omega N_h} & N_h^2 e^{j\omega N_h} \end{bmatrix}}_{\mathbf{D}_k}\underbrace{\begin{bmatrix} p_{2,k} \\ p_{1,k} \\ p_{0,k} \\ p_{0,k}^* \\ p_{1,k}^* \\ p_{2,k}^* \end{bmatrix}}_{\boldsymbol{p}_k} \tag{6}$$

Figure 2. Simplified on-board power supply net as excitation source for the battery. i_b and u_b are used for characterization.

Figure 3. The fourier transform of the current signal, obtained by exemplary simulation.

$$\bar{p}_k = 2\mathbf{D}_k^{\#}\bar{s}_k \tag{7}$$

As real valued features for regression, the absolute ($\mathrm{abs}(p_k)$) and the polar angle ($\mathrm{arg}(p_k)$) of the complex parameters p_k are used.

Table 1. Parameterization for the simulation model in Fig. 2.

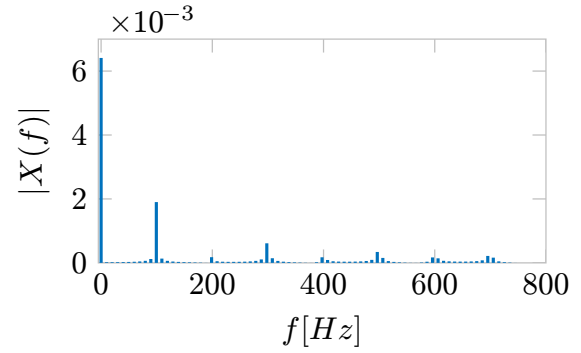| parameter | correlation coefficient | |
	abs	arg
$\bar{u}$	0.94343	
p_0	0.94014	0.16482
p_1	0.87403	0.80302
p_2	0.32694	0.28902

3 Simulation setup

To investigate the correlation of the p_k parameters with the SoC, simulations have been used as data source. The model parameters for the simulation are stated in Table 1. The battery model is a 2RC model for a commercial 18 650 lithium iron phosphate cell, parameterized by Vergossen et al. (2011). To achieve an oscillating signal, the battery has been put into a strongly simplified on-board power supply net as depicted in Fig. 2. The setup consists of a modeled battery u_b connected in series with resistor R_2, which is parallel to resistor R_1. The connection to R_1 is periodically toggled by b_{sw} with switching frequency $f = 120\,\mathrm{Hz}$. It could stand for a pulsed seat heater in parallel to a miscellaneous combination of ohmic loads.

Observing the frequency spectrum of the current characteristic in Fig. 3 resulting from that setup, it can be seen that there is a dominant component at switching frequency f. Assuming the superposition principle to be valid, this signal can be used to obtain the $p_k^{(2)}$ parameters.

The TFT used to extract the parameters was executed on windows with a length of $N = 65$ samples.

4 Suitability for SoC estimation studies

To evaluate the fitness of the parameters, first the Bravais-Pearson correlation r_{xy} is investigated. After that, tests with a polynomial fit for SoC detection are conducted.

4.1 Correlation coefficient

As an exemplary measure for the usability of parameters for regression, the Bravais-Pearson correlation coefficient r_{xy} of simulation output to the SoC is investigated.

Let $\bar{a}$ be the mean value of an arbitrary signal a, then the coefficient for signals x and y calculates as (Fahrmeir et al., 2016)

$$r_{xy} = \frac{\sum_{i=1}^{K}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{K}(x_i - \bar{x})^2 \sum_{i=1}^{K}(y_i - \bar{y})^2}} \tag{8}$$

Commonly, $0.8 \leq r_{xy}$ is perceived as strong correlation, $0.5 \leq r_{xy} < 0.8$ stands for a intermediate correlation and $r_{xy} < 0.5$ is considered as weak correlation (Fahrmeir et al., 2016). Investigating the results in that scope, it has to be kept in mind that the correlation coefficient is a measure for *linear* correlation. The results can be found in Table 2.

Following the definition from above, the terminal voltage $\bar{u}$, as well as the parameters $\mathrm{abs}(p_0)$, $\mathrm{abs}(p_1)$ and $\mathrm{arg}(p_1)$ show a strong correlation to the SoC, the remaining parameters $\mathrm{arg}(p_0)$, $\mathrm{abs}(p_2)$ and $\mathrm{arg}(p_2)$ only correlate weakly in linear sense. Though this results seem bad, its has to be mentioned that a battery is a highly nonlinear system. Therefore, polynomial dependency has also been investigated.

4.2 Polynomial regression

To get a further idea for the suitability for SoC detection of the parameters beyond linear correlation, a multivariate

Table 2. Correlation coefficient of the paramters with the SoC for an exemplary simulation.

order of	error [% SoC]			
	$e_{\mathrm{rms},1}$		$e_{\mathrm{rms},2}$	
polynomial	μ_1	σ_1	μ_2	σ_2
1	5.7188	3.8×10^{-3}	8.1122	4.2×10^{-3}
2	0.0245	2.0×10^{-5}	3.4479	2.5×10^{-3}
3	0.0091	7.2×10^{-6}	1.9777	1.6×10^{-3}
4	0.0087	0.0000	0.9610	1.0×10^{-3}

Table 3. Mean value μ_x and standard deviation σ_x of the root mean square error of polynomial regression by order of polynomial.
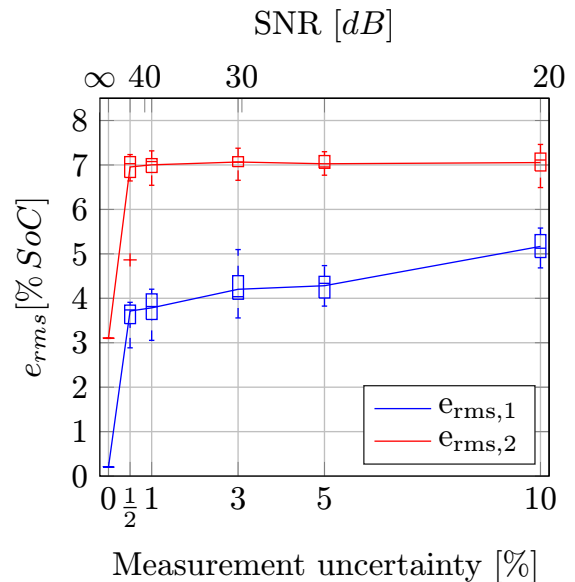
parameter	correlation coefficient	
	abs	arg
$\bar{u}$	0.94343	
p_0	0.94014	0.16482
p_1	0.87403	0.80302
p_2	0.32694	0.28902

Figure 4. SoC estimation error over measurement uncertainty and SNR, obtained by exemplary simulation.

polynomial has been fitted to the data. As fitness measure, the root mean square error ($e_{\mathrm{rms},i}$) is given for a polynomial $g_1(\bar{u}, \mathrm{abs}(p_0), \mathrm{abs}(p_1), \mathrm{arg}(p_1), \mathrm{abs}(p_2), \mathrm{arg}(p_2))$ making use of voltage measurement and the calculated parameters and $g_2(\bar{u}, \mathrm{abs}(p_0), \mathrm{arg}(p_0))$ (voltage and impedance) for comparison. The corresponding errors are $e_{\mathrm{rms},1}$ for g_1 and $e_{\mathrm{rms},2}$ for g_2. As common practice in supervised learning, the data set is split into a training set (targets) and a test set. The proportion of training to test data is $80 : 20$. The data are assigned to the sets randomly.

First, the influence by the order of the polynomial has been investigated. The mean values μ_x and standard deviations σ_x dependent on the polynomial order are presented in Table 3. To avoid influence of the set assignment process, 51 Monte Carlo runs have been conducted. To avoid inaccuracies due to rounding errors, the measurement and data matrix have been transformed into zero mean unit variance form.

It can be seen that for g_1, the mean root mean square (RMS) error is always lower than for g_2. For orders higher than 1, the mean RMS error is lower than 1 % for g_1, where for g_2 this takes a polynomial order of 4. Also, the standard deviation σ_1 is lower than σ_2 all the time, declining along with μ_1 over inclining order of the polynomial.

In the ideal case presented here, using the new features results in considerably lower deviations from the ground truth than using only the terminal voltage and the impedance. Where $e_{\mathrm{rms},2}$ only undercuts the value of 1 for a order of polynomial of 4, the error $e_{\mathrm{rms},1}$ is well below 1 % for all cases with a polynomial of order 2 or higher. The resulting tradeoff is between a high accuracy with g_1, where the com-

putational cost is higher because of the pseudo inverse on the one hand, and the computationally cheaper but less precise solution of g_2. Increasing the order of polynomial also increases the performance of g_2, but to lift it on the same level as g_1, a high number of summands would be necessary. This would also result in high computational cost, unsaid that the success of that is guaranteed.

Finally, the robustness agains the influence of noise has been investigated. The noisy simulated measurement data have been used unaltered, e.g. without filtering. As noise, usual values for measurement uncertainty have been assumed as noise variance, e.g. $e_{\mathrm{meas}} = \{0, 0.5, 1, 3, 5, 10\}\%$. In Fig. 4, $e_{\mathrm{rms},1}$ (blue line) and $e_{\mathrm{rms},2}$ (red line) are depicted over different noise levels. The graphs show the mean values over the 11 Monte Carlo runs, where the box plots display the typical statistical deviations. For g_1, the error for no noise is below 1 %, only differs little around 4 % for measurement uncertainties of 1 to 5 % and is short above 5 % with a high measurement uncertainty of 10 %. For g_2, for no noise the uncertainty is between 3 and 4 %, and stays flat at around 7 % for all other values.

For all simulations, g_1 scored a significantly lower error. The difference is around 3 percent points for low noise and only undercuts 2 percent points for a measurements uncertainty of 10 %.

5   Conclusions and outlook

In this work, the dynamic phasor and its derivatives have been presented as new impedance based features for SoC classification. Simulations based on a lithium ion battery

model have been conducted. Investigations utilizing the correlation coefficient r_{xy} showed a strong linear correlation of a subset of the features. Regression tests with a multivariate polynomial of orders above 2 yield a RMS error below 1 % SoC, which in comparison is lower than a regression featuring only terminal voltage and impedance measurements. In an exemplary experiment with a polynomial of order 2 on data with altering measurement uncertainties, the impedance features yield to a 2 to 3 percent point lower estimation RMS error.

In future work, the features could be subject to more elaborate performance testing utilizing any type of regression methods from the field of machine learning. Especially, ensembles of different regression types, potentially in combination with other SoC detection algorithms, could be with good prospects for highly precise state estimation. Furthermore, a laboratory study could be conducted to compare the results from the measurements with the ones obtained in the simulations.

Code availability. The code is subject to an unpublished dissertation, hence will not be published beforehand.

Data availability. Rights for battery model lookup tables reside with Vergossen et al. (2011) and are not provided.

References

Chang, W.-Y.: The State of Charge Estimating Methods for Battery: A Review, ISRN Applied Mathematics, 2013, 953792, https://doi.org/10.1155/2013/953792, 2013.

de la O Serna, J. A. and Rodríguez, J.: Dynamic Phasor Estimates for Power System Oscillations, IEEE Transactions on Instrumentation and Measurement, 56, 1648–1657, https://doi.org/10.1109/TIM.2007.904546, 2007.

European Parliament: Regulation (EU) No. 333/2014, European Commision, available at: http://ec.europa.eu/clima/policies/transport/vehicles/cars/ (last access: December 2016), 2014.

Fahrmeir, L., Heumann, C., Künstler, R., Pigeot, I., and Tutz, G.: Statistik, Springer-Verlag, 8. edn., https://doi.org/10.1007/978-3-662-50372-0, 2016.

Feng, F., Lu, R., Wei, G., and Zhu, C.: Online Estimation of Model Parameters and State of Charge of LiFePO$_4$ Batteries Using a Novel Open-Circuit Voltage at Various Ambient Temperatures, Energies, 8, 2950–2976, https://doi.org/10.3390/en8042950, 2015.

Keller, K. J.: Improved LiFePO$_4$ cell balancing in battery-backup systems with an Impedance Track fuel gauge, Texas Instruments Analog Application Journal, 14–17, 2013.

Klotz, D., Schönleber, M., Schmidt, J. P., and Ivers-Tiffée, E.: New approach for the calculation of impedance spectra out of time domain data, Electrochimica Acta, 56, 8763–8769, https://doi.org/10.1016/j.electacta.2011.07.096, 2011.

Piller, S., Perrin, M., and Jossen, A.: Methods for state-of-charge determination and their applications, Journal of Power Sources, 96, 113–120, 2001.

Richardson, G., Denuault, G., and Please, C. P.: Multiscale modelling and analysis of lithium-ion battery charge and discharge, Journal of Engineering Mathematics, 72, 41–72, 2012.

Schmidt, J. P., Chrobak, T., Ender, M., Illig, J., Klotz, D., and Ivers-Tiffée, E.: Studies on LiFePO$_4$ as cathode material using impedance spectroscopy, Journal of Power Sources, 196, 5342–5348, https://doi.org/10.1016/j.jpowsour.2010.09.121, 2011.

The German National Platform for Electric Mobility: Zweiter Bericht der Nationalen Platform Elektromobilität, Gemeinsame Geschäftsstelle Elektromobilität der Bundesregierung (GGEMO), available at: http://nationale-plattform-elektromobilitaet.de/die-npe/publikationen/ (last access: December 2016), 2011.

Vergossen, D., Schmidt, J. P., Ivers-Tiffée, E., Sauer, D. U., John, W., and Busché, S.: Parameterization of a Lithium-Ion Cell Model for Simulation of the Dynamic Behaviour of a 12 Volt Starter Battery, in: Kraftwerk Batterie, 2011.