Determining the State of Charge and the State of Health of a Battery Pack

Clement Limandibrata
Electrical Engineering
University of Twente
Enschede, The Netherlands
Clementlimandibrata@student.utwente.nl

***Abstract*— In this paper, a state of charge (SOC) and state of health (SOH) mobile estimation lab is developed, and the error of the setup is quantified. Multiple implementations for estimating the SOC and SOH are outlined below. However, the coulomb counting estimation method is employed to estimate the SOC, while the linear regression of the SOC is used to estimate the SOH. This is accomplished by using a current probe connected to an oscilloscope to record current levels. The data acquired inputs into a Matlab script, which calculates the SOC and SOH values. Experimental data show that this implementation can estimate the SOC value and thus the implementation can also estimate the SOH. However, a standalone coulomb counting implementation has a couple of drawbacks. These drawbacks and ways to overcome these drawbacks are detailed below. The setup is validated by the open-circuit voltage method, and it is figured out that the error of this implementation grows at $0.0013\ \frac{\Delta \mathrm{SOC}}{\mathrm{minute}}$.**

I. Introduction

The ever-increasing use of portable electronic devices in current generations, where rapid advances in technology and the decreasing prices of electronics have boosted the demand for batteries to power these portable devices. Electric vehicles (EVs), portable electronic equipment, and a variety of other power systems are examples of equipment that depend on batteries to provide power. This rising demand necessitates a better knowledge of these battery packs to ensure that they do not represent a threat. A battery may pose a threat when the limits of the batteries are exceeded, these batteries may overheat and may spontaneously combust. These batteries undergo intricate chemical and electrical processes which cause the batteries to exhibit non-linear properties. Unlike other electronic components such as resistors with linear qualities, these batteries are unpredictable due to their non-linear properties. As a result, further study is needed to better understand these batteries. When performing battery research, the SOC and SOH are two important states to be aware of. This project aims to provide a mobile SOC and SOH estimation set up to aid in future research and quantify the error of the setup.

SOC is not a directly measurable quantity such as open-circuit voltage or current from a battery pack, it must be estimated by relating some of these measurable quantities to the SOC. Several techniques have been developed to estimate SOC which will be discussed below along with their computational complexity, the accuracy of estimation, and ease of implementation. At a first glance, it may seem that the accuracy of the estimation has to be prioritized. However, these qualities are equally important for example, to achieve high accuracy of estimation, high computational complexity is required, which makes it less suitable for real-time purposes.

The techniques to estimate SOC fall under two main categories [1]. These being, the direct techniques that measure the battery's electrical parameters, which arithmetically process these values to estimate the SOC, and indirect techniques, which indirectly associates the battery's electrical parameters to the SOC, utilizing a battery model of sorts as a mediator between the battery's electrical parameters and the SOC.

A. Direct methods to estimate SOC

The first technique to estimate SOC directly is known as the open-circuit voltage (OCV) based estimation. This technique takes advantage of the relation between the OCV of the battery cell and its SOC. Since a specific OCV maps to a certain SOC value. Usually, this relation is characterized through a polynomial or a look-up table [1]. However, since this method relies on the OCV of the battery pack, high-resolution sensors are required to reliably estimate the SOC of the battery pack. In addition, the OCV needs to reach equilibrium to accurately measure it, this is what is known as battery relaxation [2]. When the discharge/charge current from the battery is removed, the OCV of the battery increases at a rapid pace at first, then gradually climbs then reaches equilibrium around two hours after the load has been disconnected [3]. As a result, it's not suited for real-time measurements.

B. Indirect methods to estimate SOC

There are a couple of distinct methods that fall under the indirect methods; a model-based implementation, an adaptive filter implementation, and a data-driven implementation [4]. These three methods use directly measurable quantities, such as the OCV of the battery pack and the battery's charging/discharging current, and use them as inputs for their respective systems.

First, the model-based techniques will be discussed. There are two main models, other models are developed from these two models. The first model is the electrical circuit model (ECM). This model estimates the battery as a network of resistors and capacitors whose values are known beforehand, either from experimental data or from data sheets and alike. This leads to a decrease in computational complexity but it can not take into account other variables such as temperature [3]. In addition, this method is an inflexible model, since the parameters of the resistors and capacitors differ for each battery and also would differ with changes in environmental factors, such as ambient temperature. The second model is the electrochemical model (EChm). This model is made up of partial differential equations (PDE) to model the battery's chemical structure and the chemical reactions that happen inside the battery. This model is highly credible, thus the estimation of SOC would be highly accurate [3]. However, the computational complexity of this model is high which makes this model unsuitable for on-board implementation. A derivative of the EChm model, the electrode-averaged model (EAM), is suitable for real-time SOC estimation. This may be attributed to its ability to handle rapid discontinuities for electrode bulk concentration, such as in EVs with restorative braking.

The second indirect method is an adaptive filter. This method combines some direct methods and a battery model. As the name suggests, the filter is a self-adjusting system, adjusting its parameters to achieve the desired outcome, this is done with the help of a feedback loop. An example of an adaptive filter is the Kalman filter (KF) which uses a set of measurements to estimate unknown variables, also known as a hidden state. A KF estimates hidden states, such as SOC or SOH, in a linear dynamic system derived from measurable quantities where the error between the expected value, from the battery model, and the observed value is minimized. Since the inner workings of a battery is a non-linear process, this filter needs to be linearized each time step with the help of a first-order Taylor expansion. This results in what is known as an extended Kalman filter. However, this filter assumes that the noise that is inherently included in the expected value, from modeling errors, and observed value, from measurement errors, is white Gaussian random processes with zero mean [1]. Which limits its field application where the noise is not guaranteed to meet the criterion above, leading to the estimation of SOC to diverge from the its true value.

Lastly, the third indirect method is adaptive artificial intelligence. This method includes methods such as Fuzzy-Logic-Based Estimation, Artificial Neural Networks-Based Estimation, and Genetic Algorithm-Based Estimation [1]. An advantage of this method is that it has a strong matching ability [4], the ability to figure out patterns from measurement values, allowing it to estimate unknown variables by figuring out patterns from training data sets (the battery model). However, this method also has the chance to over-fit the training data sets [4].

C. Methods to estimate SOH

There are no commercially available sensors to directly measure the capacitance nor the internal impedance of the batteries. Thus, the main focus to estimate SOH is to develop techniques to indirectly calculate the SOH using quantities that commercially available sensors can measure. The estimation of the SOH can be implemented in four distinct ways. A model-based, a change in SOC based, and, a differential voltage analysis/incremental capacitance analysis, and an adaptive artificial-intelligence-based estimation. These methods will be discussed further below.

There are two model-based approaches, one physics-based and the other empirical-based. The physics-based model describes the chemical reactions of the battery with the help of PDEs. These chemical reactions are closely related to the decay of the battery [4]. The physics-based model has high accuracy, however, it puts a heavy strain on computing. Which makes this model unsuitable for online applications. The other model is the empirical-based model, which is derived by fitting data from actual battery experimental data, and a battery model can be formulated. This model can achieve high computational efficiency. However, this method is labor-intensive to formulate, as substantial aging tests need to be carried out. These models also exhibit poor flexibility [4] as the model is valid for certain conditions only.

The second method to estimate the SOH is an incremental capacity analysis. This method estimates the capacity by differentiating the battery's capacity over its voltage when the battery is in constant current charging mode, $\frac{dq}{dv}$. Graphing the $\frac{dq}{dv}$ and voltage results in a graph that has peaked on it, these peaks can be used to predict the capacitance of the battery at a certain cycle. In addition, this method has a high demand for the resolution of current and voltage measurements, as a lack of resolution in the measurement of current and voltage results in quantization noise which will bias the results of the IV curve, and thus also the SOH value.

Lastly, the adaptive artificial intelligence implementation will be discussed. The methods to estimate the SOH through adaptive artificial intelligence are, but are not limited to, support vector machine and Gaussian process regression [4]. These methods have the advantage of a model-free reinforcement learning characteristic, where the model's inner workings do not need to be considered in real-time. However, these algorithms need to be ”trained” with multiple data sets, if only a few data sets are used for training, the algorithm will over-fit the training data set and would not be suitable for anything else.

As stated above, batteries are increasing in consumption and research has to be done to both make these batteries safer and better. The SOC and the SOH are two of a couple of states that are used to characterize a battery cell or battery pack. In order to quickly quantify the SOC and the SOH, this paper proposes a simple implementation to estimate both states.

The following is a breakdown of the paper's structure. The project's essential information is discussed in section II. In section III, the reader will go through how to set up the estimation of SOC and SOH and justification of choices made about the set up. In section IV, the reader can view the result of the setup. The discussion follows in section V. Finally, section VI brings the paper to a close.

II. Theoretical Framework

A. Definition

In short, the SOC is like a fuel gauge on a vehicle. it describes how much capacity a certain battery pack has left compared to the maximum capacity a certain battery pack can store. The SOC is expressed in percentages, from 1 to 0. The SOC is 1 when the battery is fully charged and close to 0 when the battery is completely discharged. A common mathematical expression that is used to describe SOC is shown equation 1 and 2, as seen in [4].

$$SOC(t) = SOC(t_0) + \int_{t}^{t_0} \frac{I(t)\eta}{Q_n} \tag{1}$$

$$SOC = \frac{C_r - C_d}{C_r} \tag{2}$$

Where $SOC(t)$ and $SOC(t_0)$ is the SOC of the battery at time t and t_0 respectively. $I(t)$ is the output current of the battery pack expressed in ampere, η is the coulombic efficiency which describes how much capacity can be discharged in comparison to how much charge is used to charge the battery pack. Q_n is the capacitance of the battery expressed in ampere-hour (Ah). C_r is the rated capacitance, and C_d is the discharged capacity, both expressed in ampere-hour (Ah).

Next, the state of health (SOH), this state is important since the battery degrades over time. It is defined as the ability of a battery to store energy in comparison to its nominal capacitance. Thus, SOH is expressed in percentages. Typically, the battery's SOH is initially, at the time of manufacture, 1. This degradation can be seen from two aspects, the battery's capacitance and its internal impedance [4]. Two mathematical expressions describing the SOH is shown in equations 3 and 4, as seen in [4].

$$SOH = \frac{C_a}{C_r} \times 100\% \tag{3}$$

$$SOH = \frac{R_a - R_r}{R_r} \times 100\% \tag{4}$$

Where C_a is the actual capacitance and C_r is the rated capacitance. R_a is the actual internal impedance and R_r is the rated impedance. A capacitance fade of 20% or an increase of internal resistance of 100% is considered to be the battery's end of life (EOL) [4].

B. Method to estimate the SOC

To estimate the SOC a method known as the coulomb counting method, which is a direct method, is employed, this method is the most popular technique due to its accuracy for short-term calculation [1]. This method requires measurements of the current output of the battery pack to estimate the change of charge in a battery pack. To do so, an equation that relates the current and the SOC is presented, this equation is equation 1. However, due to noisy measurements and rounding errors, the integral (shown in equation 1) accumulates these errors, and this method also suffers from an initial value error, on account of the lack of initial SOC value, which is usually assumed.

C. Method to estimate SOH

The method to estimate SOH takes advantage of the mathematical definition of SOC, stated in equation 1. Rearranging equation 1 results in equation 5, which is the mathematical basis for capacity fade [3].

$$\int_{t_0}^{t} I(t)\eta dt = Q_n[SOC(t) - SOC(t_0)] \tag{5}$$

Equation 5 introduces a linear relationship between the integrated current and the change of SOC, with Q_n, the actual capacitance being the slope. The actual capacity can be calculated using linear regression methods. Linear regression is an approximation of relating dependent variable(s) and independent variable(s). A line of best fit, where the distance of data points towards this line of best fit is minimized. By calculating the current capacitance, the SOH can be calculated using equation 3 with C_a being the current capacitance and C_r being the rated capacitance, shown in equation 6.

$$SOH = \frac{C_a}{C_r} = \frac{\int_{t_0}^{t} I(t)\eta dt}{SOC(t) - SOC(t_0)} / C_r \tag{6}$$

III. METHODOLOGY

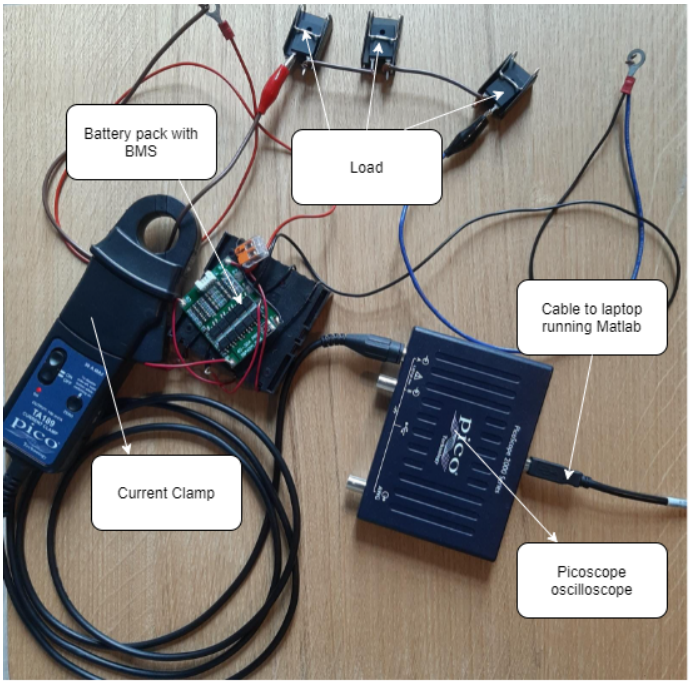

Fig. 1. Setup of the discharge circuit

A. Data collection

This implementation of SOC and SOH estimation only requires current readings as input. This current measurement was obtained using a TA-189 current clamp attached to a Picoscope 2208B oscilloscope connected to a laptop running a Matlab script. Figure 3 shows the inner workings of the Matlab script, the rhombus indicates an if statement, whose input either comes from a variable analyzed by Matlab, such as checking if there is an open connection, or input from human in the form of a message box, such as in the case of asking to plot live data. The rounded rectangles are processes that the Matlab script executes.

A single battery pack consisting of four ICR 18650-26J connected in series was used as a sample battery pack, whose specifications are listed in table I. These battery cells are not directly connected to the load, to increase the safety and the longevity of the batteries, a battery management system (BMS) was connected between the load and the battery pack. A BMS can increase the safety and the longevity of the batteries by acting as a fuse when certain thresholds are met, such as overcharged voltage or the maximum current drawn.

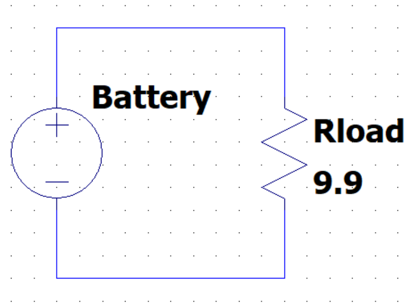

Fig. 2. Discharge circuit

As a load, a power resistor, with a resistance of 9.9 Ω, as seen in figure 2. This resistor is made up of three 3.3 Ω power resistors connected in series, as seen in figure 1 labeled as load. The reason for a low ohmic resistor was so that the current output was high, minimizing random errors, which arise from external factors. In addition, the battery was fully charged before discharging. This is because the Matlab script used to calculate the SOC and SOH assumes the starting SOC is one, or completely charged. This is because the starting point SOC can be guaranteed more easily when it is full, compared to an intermediate value between 1 and 0. The battery is always charged past the constant current phase and constant voltage phase, until the current towards the battery is 0.05 A and the voltage between the positive and negative terminals of the battery pack is 16.8 V. This is to compensate for the lack of a starting point and to standardize the amount of charge that is in the battery pack.

TABLE I
SPECIFICATION OF BATTERY

Type	ICR 18650-26JM
Nominal Capacitance/mAh	2600
Nominal Voltage/V	3.63
Charged Voltage/V	4.2
Maximum discharge current/A	5.200

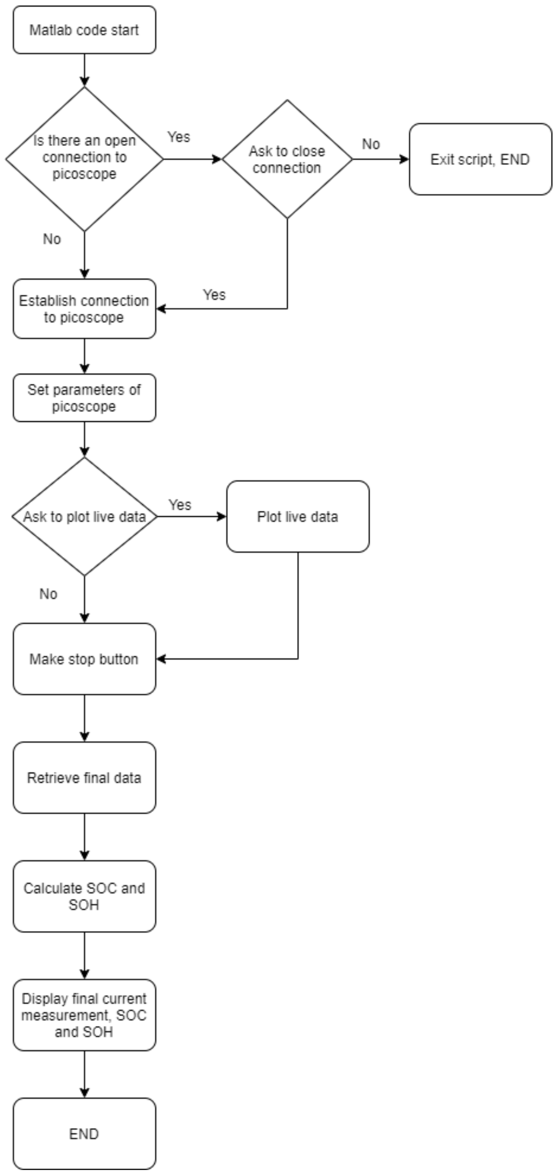

Fig. 3. Flow chart of matlab script

B. Data processing

As previously indicated, the Matlab script handled all of the data processing. As there was high-frequency noise present in the readings, of the current clamp a low pass filter was applied to the raw data from the current clamp. This noise is present due to the cable acting as an antenna, picking up background transmissions and incorporating them into the current measurements. In addition, the current clamp records current as a voltage with a conversion of $100\,\frac{\mathrm{mV}}{\mathrm{A}}$, thus the script has to account for this conversion.

1) SOC estimation: For the calculation of SOC, equation 1 was implemented. The integration of current was done with the help of the Matlab function cumtrapz, which provides intermediate values for graphing purposes. However, this function assumes that the distance between each successive time points to be one second, thus a spacing increment needs to be multiplied to the result of the cumtrapz. This spacing increment is equal to $\frac{B-A}{N}$ where 'N' is the number of samples, 'A' is the starting time and 'B' is the end time. The starting SOC ($SOC(t_0)$) was set to 1, so the battery pack was assumed to be full at the start of the Matlab script when the current measurement started. Lastly, the capacity (Q_n) was the rated capacitance of the battery.

2) SOH estimation: The linear regression was implemented with a polynomial regression tool available in Matlab, polyfit and polyval. Polyfit provides the polynomial coefficient of the fitted SOC curve, and polyval evaluates the polynomial created by polyfit at each time point. The SOH estimation takes into account the whole experiment to reduce random errors and errors due to fluctuation of current measurements, so the SOC at the initial time (t_0) was the SOC value at the start of the first current measurement and the SOC at time t was the end of the experiment ($SOC(t)$). Following that, equation 5 was used to calculate the current capacitance and lastly, equation 3 was used to calculate the SOH. The SOH was displayed in a popup message created with the msgbox function in Matlab.

C. Performance evaluation

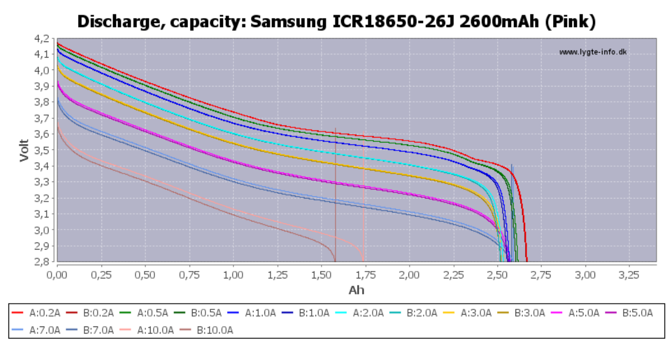

Fig. 4. OCV-discharge capacity graph [5]

The performance of the SOC estimation has been validated by checking the OCV of each battery cell and calculating how much capacitance has been discharged according to the OCV-discharge capacity graph shown in figure 4 and equation 3. This calculated SOC has been regarded as the true value, y, and the SOC from the Matlab script was the estimated value, $\hat{y}$. The error rate was calculated with the mean absolute error (MAE), shown in equation 7. Which calculates the absolute distance each measured value was from the true value. Moreover, another metric, the mean absolute percentage error (MAPE), was included. This metric shows MAE as a percentage of the true value, to give more meaning to the error metric since the MAE is a deceptively low value, a value below 1.

$$MAE = \frac{1}{n}\sum_{n=1}^{n}|y-\hat{y}| \tag{7}$$

$$MAPE = \frac{1}{n}\sum_{n=1}^{n}\frac{|y-\hat{y}|}{y} \tag{8}$$

Equation 7 is as seen in [4] and equation 8 is as seen in [6]. In addition, since the coulomb counting is also prone to cumulative error, the error rate has also been calculated with linear regression, of the first order, based on the difference of the true value and the estimated value of the SOC. The reason for using a first-order linear regression method is because the background noise is assumed to have a non-zero mean, thus the noise can be assumed to be of constant magnitude. As a result of the integration of noise into the SOC graph, the error was assumed to be additive, since the integration of a constant magnitude results in a graph whose value increases at a constant rate.

IV. Results

A. SOC estimation

The output of the SOC estimation Matlab algorithm along with the final current measurement is shown in figure 5. The Matlab algorithm also works when the battery pack is being charged too, as seen in figure 6.

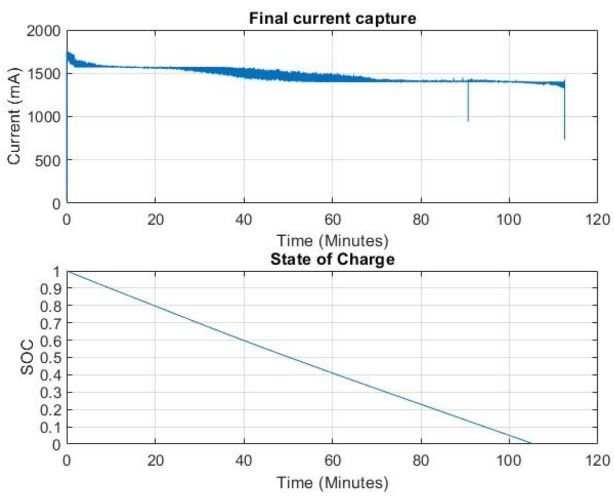

Fig. 5. Current and SOC graph, discharging at 0.6C

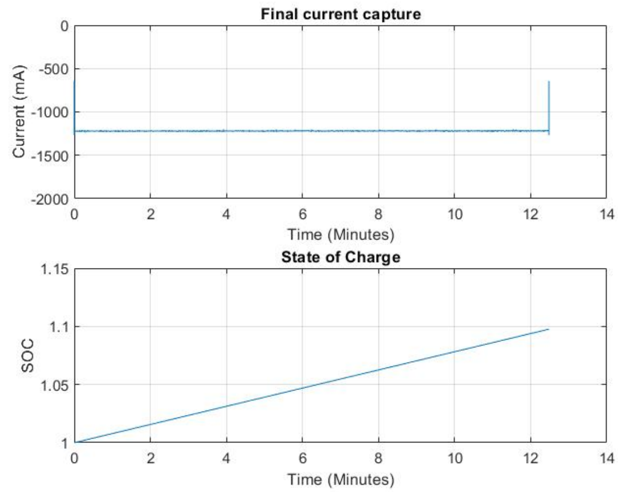

Fig. 6. Current and SOC graph, charging at 0.5C

The battery pack could only be discharged for a maximum of 112 minutes, at 0.6C, before the BMS trips, and the voltage between the P+ and the P- terminal goes to zero, and the current drops too. As can be seen in figure 7, which is the voltage reading of an individual cell at the last 18 minutes which has lasted 112 minutes, the cell's voltage declines rapidly until it reaches 2.53 V, at which point it recoils.

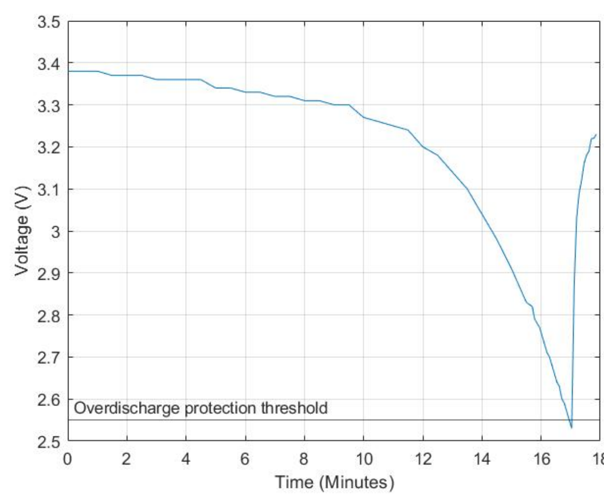

Fig. 7. Voltage of Cell

B. SOH estimation

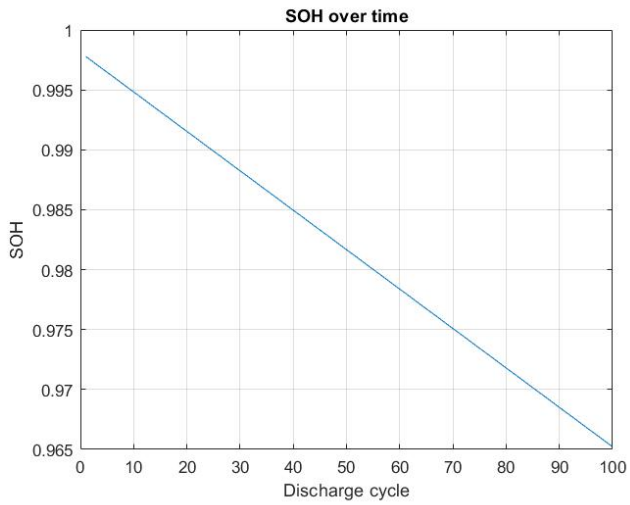

Fig. 8. Degradation of SOH

The SOH estimation was estimated and graphed against the number of discharge cycles, as seen in figure 8. This line is derived from obtaining a line of best fit between 10 discharge cycles and extrapolated to predict the SOH after 100 discharge cycles. It was found that the SOH degraded at a rate of 0.000329 per discharge cycle.

C. Performance evaluation

As stated in the methodology section, the accuracy of the coulomb counting implementation was validated by the OCV method described above. In addition, the OCV-discharge capacitance curve depends on the current drawn from the battery pack as seen in figure 4. Given that the experiments were done in the same conditions and discharged on the same resistor, as seen in figure 2, each experiment is comparable to each other. The average current drawn from the battery pack is $1.56\,\mathrm{A}$. The absolute difference between the true value of SOC, from the OCV method, and the measured value of SOC, from the Matlab script, was graphed against the duration of the measurements. Next, a linear regression method of the first order was applied to those points and was plotted on the same graph as seen in figure 9. From figure 9, the error rate was calculated to be $0.0013\,/\mathrm{minute}$ in other words, the measured SOC deviates from the true SOC value by $0.0013\,\frac{\Delta\mathrm{SOC}}{\mathrm{minute}}$ of experimental duration. In addition, the graph shows that there is an initial error of $0.0013\,\Delta\mathrm{SOC}$.

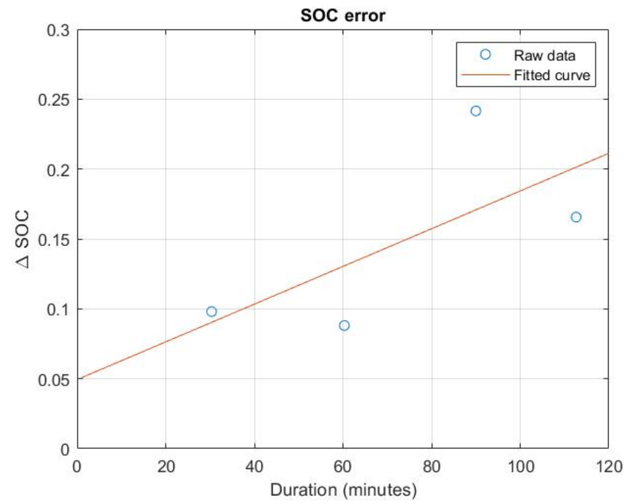

Fig. 9. Difference in predicted and measured SOC vs duration of measurement

With equation 7 the MAE was calculated to be 0.148 and the MAPE was calculated to be 0.630. However, omitting the trial with the longest duration results in an MAE of 0.143 and a MAPE of 0.308.

V. Discussion

The setup proposed is working as intended. However, the error rate is unacceptable, since the error accumulates and may grow indefinitely. Methods for overcoming the constraints of the implementation provided in this study are described in this section. A main limitation of the standalone coulomb counting method is that this method is highly error-prone from several sources. One is the accumulation of measurement error, from rounding errors, and noise, from the wires, acting as an antenna and receiving radiation from external electronic equipment. These errors accumulate in the integration of current (in equation 1) which leads to an increase in error rate with a longer duration of the current measurement. A further source of error is the starting point of the SOC estimation, which is assumed to be full at the start of the Matlab script. This is due to the setup's limitation of not being able to estimate the initial SOC. This is due to fact that current measurement alone can only figure out the change in SOC but not a specific SOC value at a certain time. In addition, the Matlab script has no way to check if the battery pack is fully charged initially. This discrepancy reflects on the successive estimation of SOC. These limitations may be overcome by employing the OCV method to both determine the initial SOC of the battery pack and reinitialize the SOC when the expected error crosses 1% with the help of figure 9.

To apply the OCV method, a voltage reading has to be obtained, to measure the OCV of the battery pack. This reading can be implemented using a voltmeter attached in parallel to the battery pack, thus reading the voltage over the whole battery. The OCV of a single battery can be calculated by dividing the OCV of the battery pack by the number of battery cells in the battery pack. In addition, the figure 4 should be characterized in a lookup table. First, the OCV of the battery pack will be measured when the Matlab code is run and from that, the starting SOC may be initialized. Next, after the error increases to 1%, from the rate of figure 9 it occurs after 6.7 minutes, the OCV will be measured again and the SOC is calculated and reinitialized.

VI. Conclusion

In this paper, an onsite SOC and SOH estimation was implemented with the coulomb counting method and linear regression method, respectively. The method proposed can successfully capture the nonlinear interaction between the SOC and SOH, and measurable quantities, such as current, according to test findings. The resulting setup results align with previous findings, in terms of the errors encountered along with its ability to relate current reading to the actual SOC value. However, along with these test findings, the shortcomings of a stand-alone coulomb counting method are shown. One of these shortcomings is the cumulative error, which is an error that increases with sample size, this can be seen in figure 9 where the line of best fit, fitted curve, shows that the difference of true and measured SOC and the duration of the measurement is directly proportional. This demonstrates that with an increase of measurement duration, the distance between the true SOC and the measured SOC gets further, thus increasing the error. Furthermore, omitting trials with higher duration of measurement reading reduces the value of MAE and MAPE. This demonstrates that measurements with a longer time stray from their real value more. Besides this point, the restriction of obtaining data past 112 minutes was caused by the BMS' over-discharge protection, as can be seen in figure 7, the voltage of the cell goes below the over-discharge voltage threshold which is 2.55 volts per cell.

References

[1] J. Rivera-Barrera, N. Muñoz-Galeano, and H. Sarmiento-Maldonado, "SoC Estimation for Lithium-ion Batteries: Review and Future Challenges," *Electronics*, vol. 6, no. 4, p. 102, Nov. 2017. [Online]. Available: http://www.mdpi.com/2079-9292/6/4/102

[2] A. Li, S. Pelissier, P. Venet, and P. Gyan, "Fast Characterization Method for Modeling Battery Relaxation Voltage," *Batteries*, vol. 2, no. 2, p. 7, Jun. 2016, number: 2 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2313-0105/2/2/7

[3] Z. Xi, R. Jing, C. Lee, and M. Hayrapetyan, "Recent research on battery diagnostics, prognostics, and uncertainty management," *Advances in Battery Manufacturing, Service, and Management Systems*, p. 175–216, 2016.

[4] X. Hu, F. Feng, K. Liu, L. Zhang, J. Xie, and B. Liu, "State estimation for advanced battery management: Key challenges and future trends," *Renewable and Sustainable Energy Reviews*, vol. 114, p. 109334, Oct. 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1364032119305428

[5] "Test of Samsung ICR18650-26J 2600mAh (Pink)." [Online]. Available: https://lygte-info.dk/review/batteries2012/Samsung%20ICR18650-26J%202600mAh%20%28Pink%29%20UK.html

[6] A. de Myttenaere, B. Golden, B. Le Grand, and F. Rossi, "Mean Absolute Percentage Error for regression models," *Neurocomputing*, vol. 192, pp. 38–48, Jun. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231216003325