Politecnico di Torino

Master thesis in automotive engineering

Battery state of health and state of charge estimation with related application

Candidate:

Liu hengrui

Han zhuoyang

Supervisor:

Prof. Stefano CARABELLI

Abstract

With the growing electric vehicle market, it has significant energy saving and emission reduction compared to the traditional internal combustion engine. Since the new power and energy supply system is different from the traditional internal combustion engine, which also brings many new challenges to the OEMs.

This thesis is mainly aimed at the optimization problem induced by the increasingly large electric vehicle market. Due to the unique characteristics of electric vehicles and traditional internal combustion engine vehicles, such as the different power source-battery, many new research directions and challenges have been brought.

In order to achieve real-time monitoring of battery health management and residual battery power during driving, we need to introduce two important parameters from the battery management system: SOC (state of charge) and SOH (state of health). Due to the battery attenuation mechanism and complex working environment, accurate real-time prediction of battery status has always been a difficult problem for manufacturers to face.

The real world's problem is that Knowing SoC and SoH is extremely important for how to use the battery correctly, but due to the complex chemical reactions inside the battery, there's no sensor can do this directly.

Therefore, we introduced a prediction method based on extended Kalman filter of MATLAB and SIMULINK environments and carried out a series of tests and simulations to verify the real battery parameters provided by battery manufacturers.

In short words, the thesis aimed to design and validate an algorithm to estimate states of the battery, which is the core function of Battery Management System.

Table of contents

Chapter 1

Introduction

Electric cars are very popular today, considering their eco-friendly, zero emission from tank to wheel point of view and better NVH during driving, the electrification is the trend of the future, however, there are still some big challenges and technical issues, one of them is to provide a precise estimation of the state of charge (SoC) and state of health (SoH), and this will be the focus of this article.

The simplest way we used to estimate the state of charge is called coulomb counting: integral the current out of the battery and divided it by the maximum capacity. This way is already widely applied nowadays, but it has a significantly drawback: without the feedback, the error(in many case is the noise of sensor) will be accumulate and lead to a unreliable result, also in this case the accuracy is strongly depends on the initial condition, but in real case we normally don't have so much initial conditions parameters to support this approach.

The solution could be adding the feedback, so the system has the ability to turnback from the wrong position and follow the right approach, in our case the solution is the Extended Kalman filter.

The document is organized in the following way:

- Chapter 2 gradually introduces to the main topics with a top-down approach.
- Chapter 3 describes the most relevant estimation approaches that exist in the literature
- Chapter 4 presents the work and the adopted models
- Chapter 5 and Chapter 6 present the work with respect to the proposed estimation schemes

Chapter 2

An overview on the main topics

From internal combustion engine switch to an electric vehicle, aimed to solve several problems but also introduce new challenges

One of the main reasons is related to the air pollution: air pollution is the effect caused by concentration of solids, liquids or gases in the air that have a negative impact on the earth environment and people.

The pollution from road transport is typically: CO, HC, NOx, PM, and CO2 which gives a significantly impact on greenhouse phenomenon, increasing sales of

electric cars and improve vehicle efficiency could be an effective way to reduce the related pollutions, even though, from an ecological perspective, the electric vehicle production is still a problem, their usage allows to eliminate air pollution problem and can contribute to the realization of a real transportation revolution.
In order to remain competitive with conventional fossil fuel vehicles, electric vehicles must have (as much as possible) enough range, fast charging, and enough charging stations, thanks to the rapid development of the battery technologies and laws promulgated by local governments, we are getting closer to these targets, electric vehicles can even surpass traditional fuel vehicles in some aspects, such as comfort, safety and autonomous driving due to electronic control, this article will focus on the battery technology part.
In order to properly manage the operation of the electric vehicle, a battery management system (BMS) has to be introduced, the detailed functions f it will be present in the next paragraph, but the core function is to estimate the stage of charge (SoC) and the state of health (SoH), an accurate measurement of these two states is critical for battery use: in electric vehicles SoC represent autonomous of the vehicle and the life of an electric vehicle is largely determined by the battery, that is the SoH.
For this reason, SoC and SoH estimation problem has a remarkable industrial value and has become a hot research topic in the last decade.

In the next paragraphs the main aspects regarding the BMS and the battery are detailed explanation, and in paragraph 2.5, the fundamental parameters of interest, namely the SoC and the SOH, are introduced.

2.1 Battery Management System

A battery management system (BMS) monitors manages the advanced features of a battery, which is an embedded system inside the vehicle electronic system, ensuring that the battery operates within its safety margins. The BMS serves as the brain of a battery pack, some typical tasks that BMS must perform:

1. Measurement the current, voltage, temperature of the battery continuously.

2. Estimate the stage of the battery (SoC, SoH)

3. Charge and discharge control

4. Safety management and diagnosis (overtemperature protection Over discharge and overcharge protection)

5. Balance between different cells inside battery pack, specially in the case of active balancing: Transfer excess power to high-capacity cells during active equalization charging, and transfer excess power to low-capacity cells during discharge

Nowadays the most widely used battery in electric vehicles are lithium-ion batteries, depend on the cathode materials it can be NMC, which uses lithium, manganese, and cobalt oxide as cathode material while LFP uses a lithium iron phosphate chemistry, the difficulty lies in how to estimate the SOC and how to control and adjust the charge-discharge interval to avoid rapid battery decay.

In terms of topology, BMS can be divided into two types: Centralized and Distributed according to different project requirements.

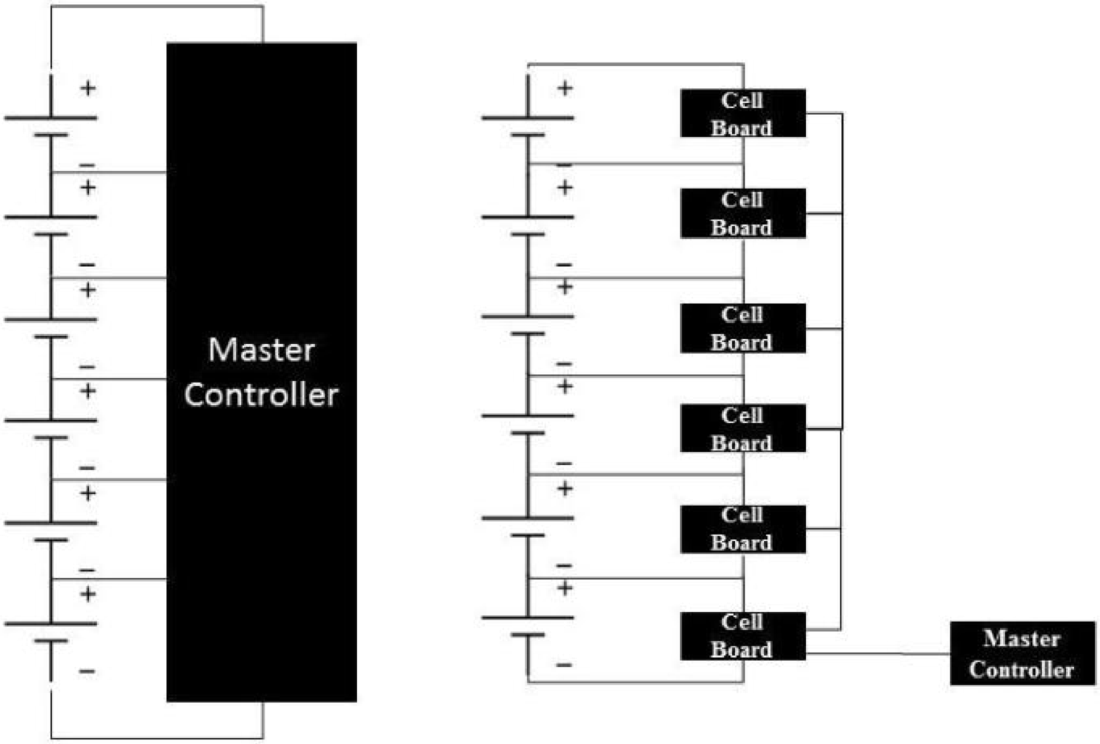


Figure 2.1: centralized (left) and distributed(right) topologies of BMS

Centralized BMS has the advantages of low cost, compact structure, and high reliability. It is generally used in scenarios with low capacity, low total pressure, and small battery system volume, such as power tools, robots (handling robots, power-assisting robots), IOT smart homes (sweeping robots, electric vacuum cleaners), electric forklifts, electric low-speed vehicles (electric bicycles, electric motorcycles, electric sightseeing cars, electric patrol cars, electric golf carts, etc.), mild hybrid vehicles.

With the continuous development of passenger car power battery systems to high capacity, high total voltage and large volume, the distributed architecture BMS is mainly used in plug-in hybrid and pure electric vehicles, the distributed BMS architecture can better realize the hierarchical management of module level (Module) and system level (Pack). The slave control unit CSC is responsible for voltage detection, temperature detection, balance management (some will independently separate the CSU unit) and corresponding diagnostic work for the monomers in the Module; the high voltage management unit (HVU) is responsible for the total battery voltage of the Pack. , bus total voltage, insulation resistance and other states are monitored (the bus current can be collected by Hall sensors or shunts); and CSC and HVU send the analyzed data to the main control unit BMU (Battery Management Unit), and BMU performs battery System BSE (Battery State Estimate) evaluation, electrical system state detection, contactor management, thermal management, operation management, charging management, diagnostic management, and implementation of internal and external communication network management

The below figure illustrates how the main elements interact to each other by depicting a traditional BMS conceptual schema:

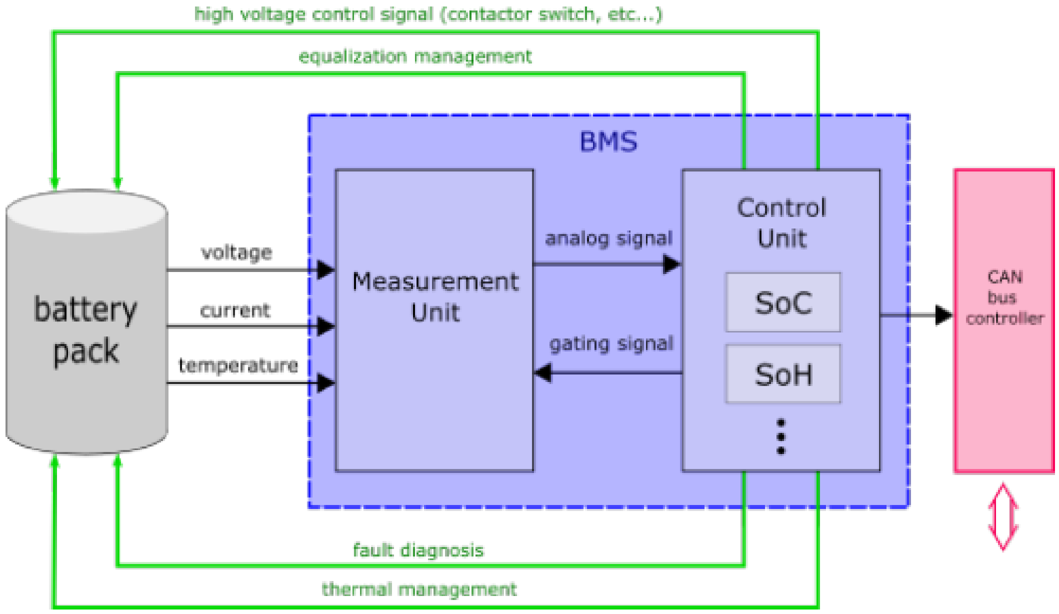


Figure 2.1: Example of a typical BMS conceptual schema [1]

The principal component of the BMS is the control unit, which runs a specific algorithm to estimate the state and ensure the safety of the battery, the inputs of the control unit are typically voltage, current and temperature, which are handled by a measurement unit, finally interact with other components inside the vehicle by use CAN bus.

2.2 composition of the battery pack

The battery pack, together with it's cover and auxiliary components , consist of the energy storage component of the electric vehicle, moving from lead to lithium solutions, in order to have high power and energy levels at the same time, a typically solution is to connect lots of modules parallel connected moreover, a certain number of cells inside the module are series connected, , that represent the very atomic energy elements of the battery. Modules embed some electronics that help the BMS to monitor the state of the cells by acquiring measures of voltage, current and temperature thanks to the presence of corresponding sensors, the benefit of this kind of integration solution not only for the performance and efficiency, but also increasing the importance of the electrical ,thermal and mechanical management (from battery cells to battery modules and battery system).

As a direct effect, the on-board integration is a key issue and volume, weight and costs have to be considered at battery system level, and on the operation lifetime/calendar life, in the decade, the most widely applied solution is cell-module-pack as mentioned before, nowadays a newest technology called cell-to-pack (CTP) is being developed, in which the modulo is replaced by the vehicle underbody structure itself, and the upper housing protection plate become the vehicle floor, since it's an immature technology until today, this article will only introduce a traditional battery pack structure.

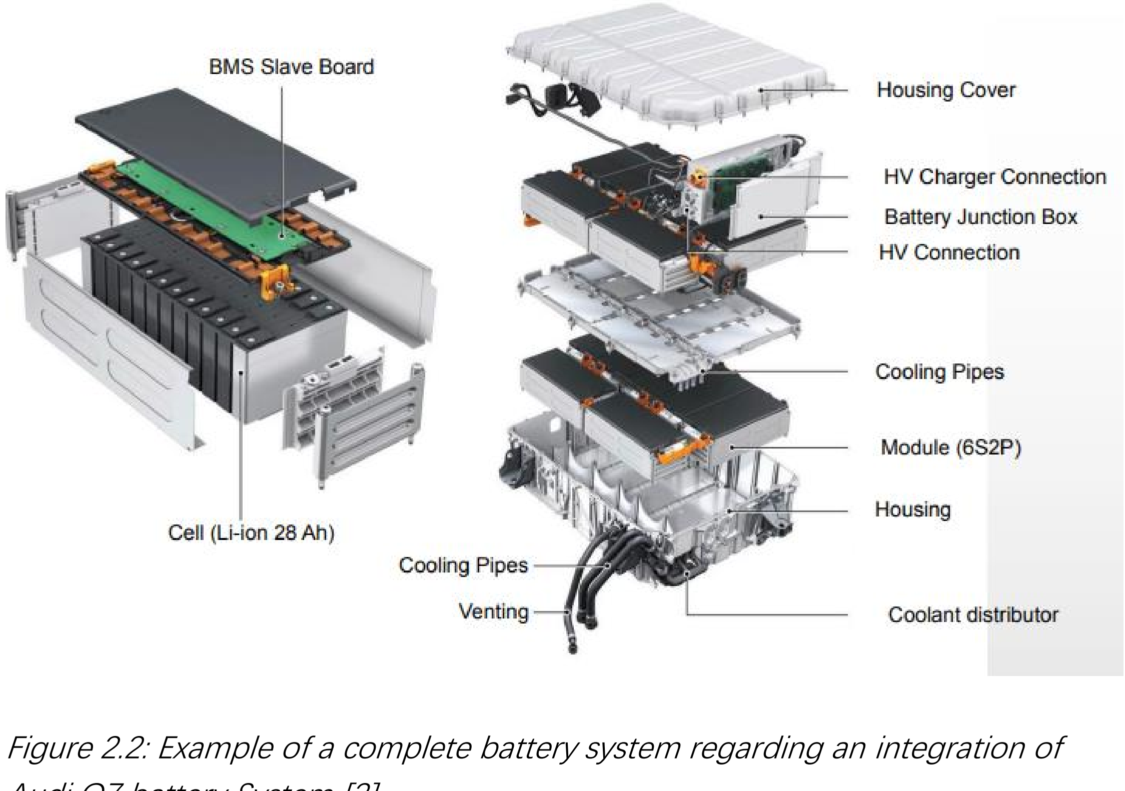
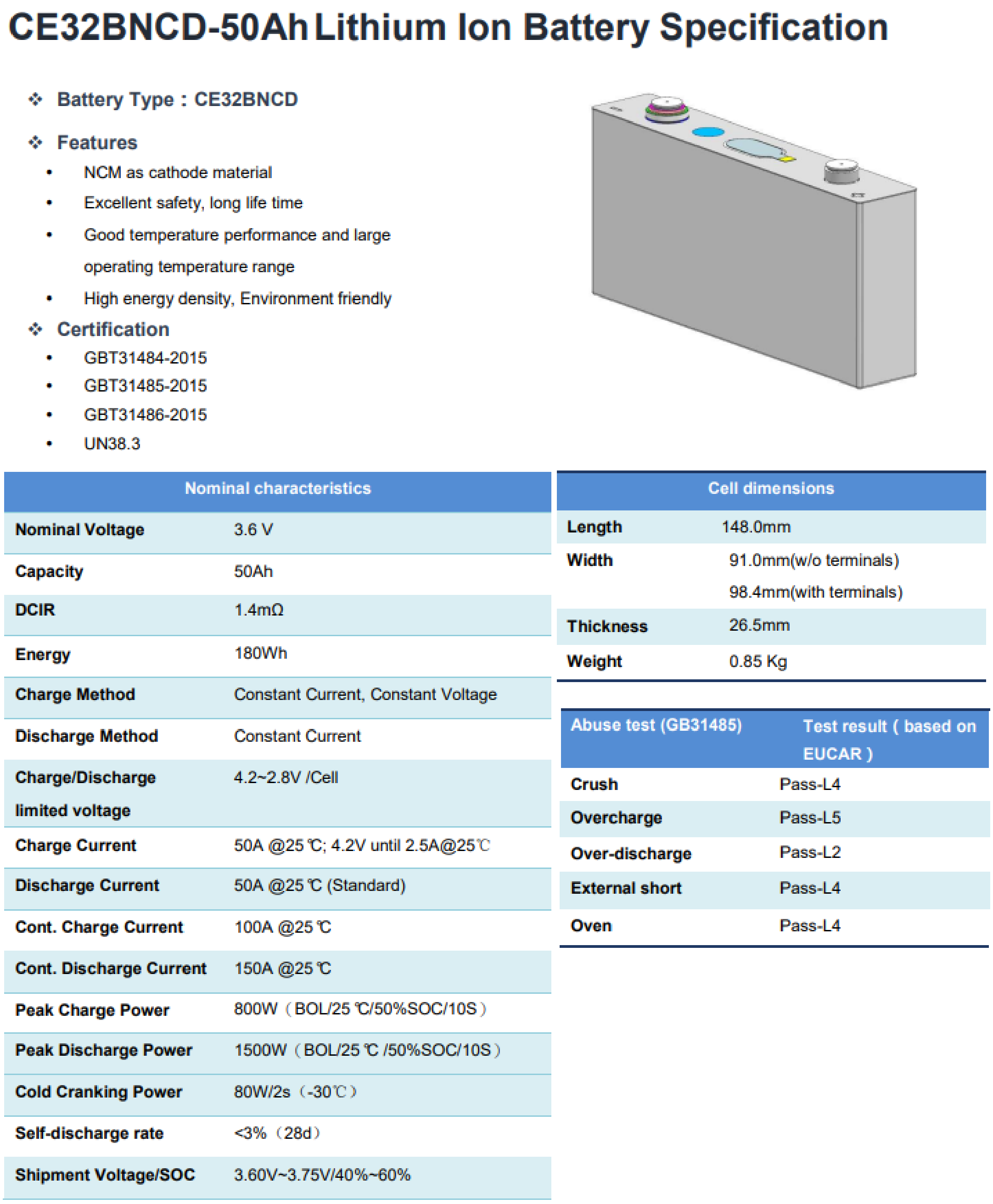


Figure 2.2: Example of a complete battery system regarding an integration of Audi Q7 battery System [2]

Nowadays, lithium-ion batteries are considered the leading battery typology for their attractive properties thanks to their important characteristics :high-density energy, long-life span a low production cost.

In the most common case, the battery manufacturer must be able to provide the detailed information of the cell, for example: nominal capacity, nominal voltage, max current, max voltage, safety operating temperature range …… this kind of information can be given by the table or graph form, for the future application and analysis



CE32BNCD-50Ah Lithium Ion Battery Specification

❖ **Battery Type：CE32BNCD**

❖ **Features**

- NCM as cathode material
- Excellent safety, long life time
- Good temperature performance and large operating temperature range
- High energy density, Environment friendly

❖ **Certification**

- GBT31484-2015
- GBT31485-2015
- GBT31486-2015
- UN38.3

Nominal characteristics	
Nominal Voltage	3.6 V
Capacity	50Ah
DCIR	1.4mΩ
Energy	180Wh
Charge Method	Constant Current, Constant Voltage
Discharge Method	Constant Current
Charge/Discharge limited voltage	4.2~2.8V /Cell
Charge Current	50A @25℃; 4.2V until 2.5A@25℃
Discharge Current	50A @25℃ (Standard)
Cont. Charge Current	100A @25℃
Cont. Discharge Current	150A @25℃
Peak Charge Power	800W（BOL/25℃/50%SOC/10S）
Peak Discharge Power	1500W（BOL/25℃ /50%SOC/10S）
Cold Cranking Power	80W/2s（-30℃）
Self-discharge rate	<3%（28d）
Shipment Voltage/SOC	3.60V~3.75V/40%~60%

Cell dimensions	
Length	148.0mm

Width	91.0mm(w/o terminals)
	98.4mm(with terminals)

Thickness	26.5mm
Weight	0.85 Kg

Abuse test (GB31485)	Test result (based on EUCAR)
Crush	Pass-L4
Overcharge	Pass-L5
Over-discharge	Pass-L2
External short	Pass-L4
Oven	Pass-L4

Figure 2.3: A screenshot about a section of a battery cell datasheet illustrating the main nominal characteristics of a CE32BNCD lithium-ion battery cell. [3]

2.3 Electrochemical reactions inside battery cell

Batteries are electrochemical energy storage devices which able to convert the chemical energy contained in their active materials directly into electric energy: the conversion process occurs through an electrochemical oxidation-reduction(redox) which in most case reversible.

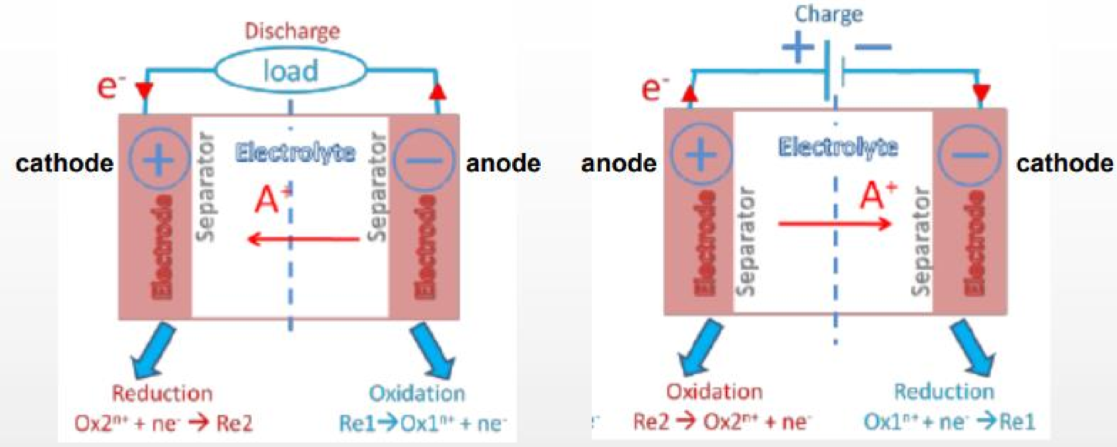


Figure2.3: an example of an electrochemical cell undergoing discharging(left) and charge(right) [4]

As the figure 2.3 shown, the charge collector(electrode) where the oxidation reaction (loosing electrons) takes place is called anode, meanwhile the reduction reaction (acquiring electrons) takes place is called cathode, the whole battery electrochemical process is the sum of one reduction reaction and one oxidation reaction.

Moving to the Lithium-ion batteries, the first commercial li-ion battery was produces by Sony in 1991 [5], this kind of battery employ lithium storage compounds as the positive and negative electrodes, the lithium ions ‘rock’ back and forth between the positive and negative electrodes as the cell is charged and discharged

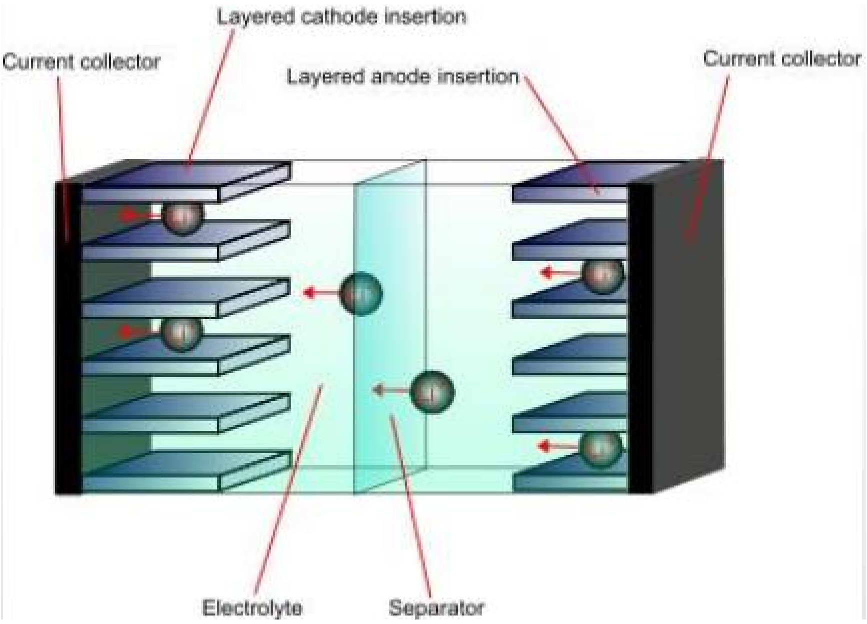


Figure 2.3.1: Lithium-ion battery during discharging [4]

The battery cells, which is the basic unit as explained in 2.2, was made of:

Anode and cathode electrodes: negative and positive electrically conductive current collectors and plates with coated catalyzer substrates.

Separator: porous membrane between the positive and negative electrodes able to insulate the electrodes each other and make pass only positive ions.

Electrolyte: in most case liquid, acting or as a carrier for ion flow between the electrodes or as an active participant in the electrochemical reaction

Current collectors: one positive and one negative, connect the cell to the other cells or electric sources.

Packaging: mechanical part used to cover and insulating the cells.

2.4 Charging/discharging operations

A main challenge for the electric vehicle nowadays is the time for the battery charging, typically an electric vehicle needs half hour to one hour to fully charge, correspondingly, traditional fossil fuel vehicles only need a few minutes, it means A lot of infrastructure needs (charging station) to be built and big wave time is wasted.

In theory, the energy enters the battery during charge is the product of power and time, and the power is the product of current and voltage, it was described with C rate, for example 1 C rate means the current during charge corresponding the nominal current of the battery, and needs one hour to fully charge (or discharge), 2 C rate means the amplitude of charging current is double and the time needed is half.

But shortening the charging time cannot be determined simply by increasing the amplitude of charging current, with standard Li-ion cell technologies, high C rate in charging can cause severe negative effects cause of the Joule effect, that can dramatically reduce the battery available capacity and reduce the battery life, some technology aimed to reduce the charging time will be present in this photograph.

Constant current – constant voltage, the most used and safety way applying on the electric vehicle charge, also take the advantage from the cost view, this kind of charging profile is divided in two part: constant current part, which using high C rate aimed to charge the battery as soon as possible, once up to a threshold temperature based voltage is reached, constant voltage part begin, charging current in this part progressively decreases, which means need long charging time but aimed to fully charge the battery

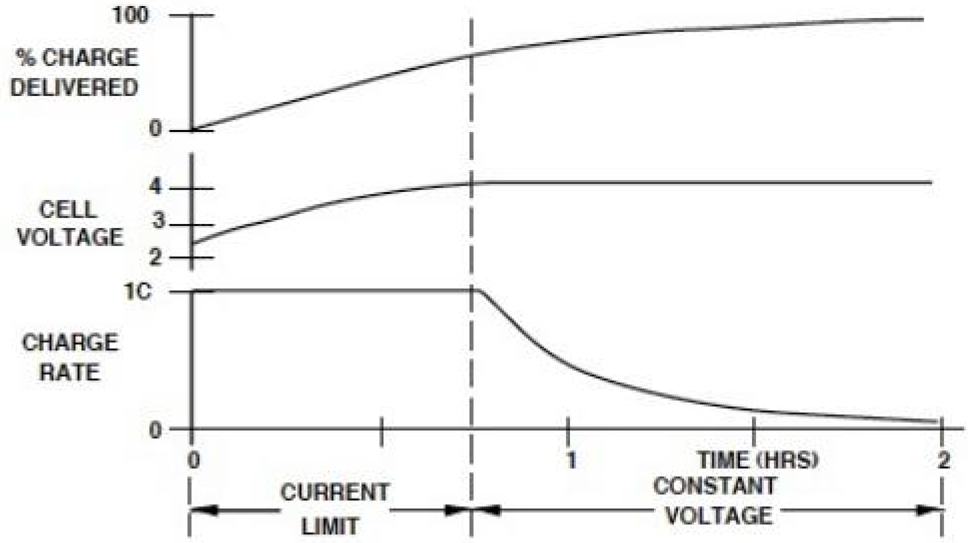


Figure2.4.1: CC-CV charging profile

Pulse charging algorithm, from the paper results [6]

, when properly designed and implemented, can result in significant benefits in terms of improved battery charge and energy efficiencies, longer battery cycle life, and decreased battery charge time. Pulse charging involves the use of charge current pulses, which have carefully controlled duty cycle, from the battery's equivalent circuit view, when the high-frequency current passes through, the capacitor of the circuit is short-circuited, and the current does not flow through the resistance of the circuit where the capacitor is located, which greatly improves the charging efficiency, as the figure below shown.

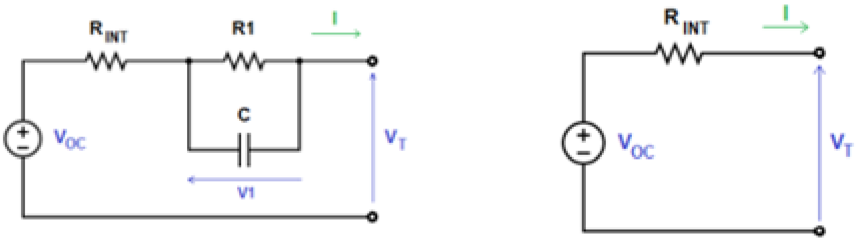


Figure 2.4.2: when a high frequency current flow, the equivalent electric circuit of battery (left) become right

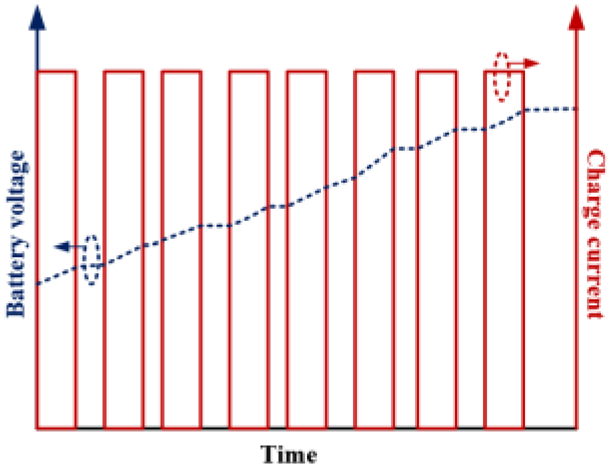


Figure 2.4.3: . Waveforms of a pulse charger under operation [6]

Aimed to reach the highest possible SoC in the lowest possible time, limiting the acceleration ageing effects, the pulse charge algorithm needs to properly be acting on the control parameter: pulse steps frequency, pulse stage on relaxation stage ration, current C rate.

A comparison has made of the CC-CV and pulse charging shown the benefit on the charging time and battery life:

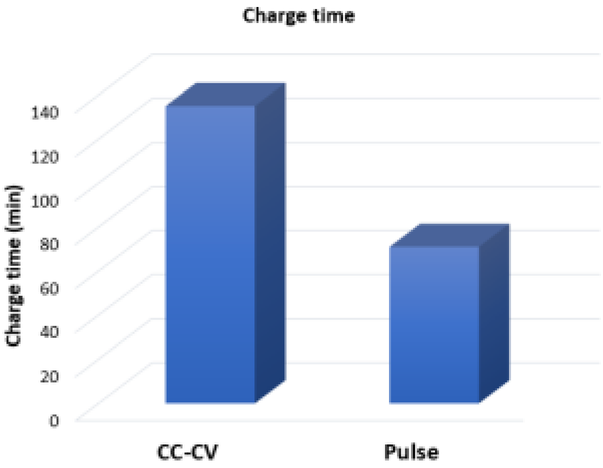


Figure 6. Comparison of charge time between CC–CV and pulse charged batteries. [6]

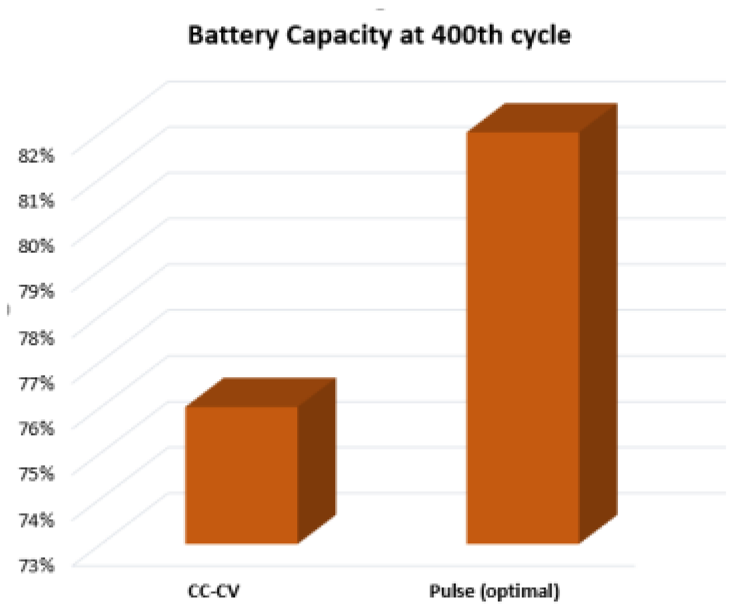


Figure 7. Remaining battery capacity after 400 charge/discharge cycles at room temperature for standardly (CC-CV) charged and optimally pulsed charged LiPo batteries. [6]

2.5 State of Charge and State of Health

As mentioned in paragraph 2.1, the core function that BMS has to be provide is to estimate the State of charge (SoC) and State of health (SoH)
the definition of SoC of a battery is the residual capacity divided by the nominal capacity which corresponding the actual SoH of the battery

$$SoC(t) = \frac{Q(t)}{Q_n} \times 100\%, \quad t \geq 0$$

The residual capacity Q(t) [A·h] corresponds to the amount of charge that battery able to discharge, meanwhile the nominal capacity Qn [A·h] corresponds to the total amount of charge that the battery in a fully charged condition, the definition of DoD can be directly extended from the definition of soc, which is the amount of charge that already dissipated of the battery:

$$DoD = 1 - SoC$$

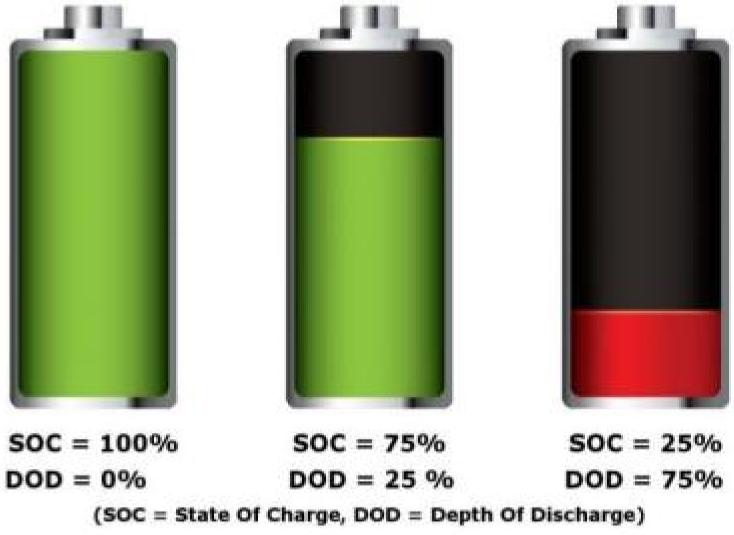


Figure 2.5: Conceptual representation of SoC and DoD [7].

In a traditional internal combustion vehicle , the remain amount of fuel can be easily measured by applying a sensor in the fuel tank. In an electric vehicle, which is no possible to direct measure the remain capacity of the battery, but it has be estimated from the measurable signal: current, battery terminal voltage, average temperature of the battery and so on, that’s why providing an accurate estimate for capacity remains (SoC) is still a big challenge nowadays.

The aging of the LI-ion battery is a long-term gradual process, and the health status of the battery is affected by various factors such as temperature, which is generally considered to be the main factor affecting the health of the battery. Temperature has a dual effect on the performance of the battery. On the one hand, high temperature will speed up the chemical reaction rate inside the battery, improving the efficiency and performance of the battery, and at the same time, high temperature will also accelerate some irreversible chemical reactions. Occurs, resulting in the reduction of the active material of the battery, causing the aging and capacity decay of the battery. Some experimental data show that high temperature will accelerate the growth of the SEI film of the battery electrode, and the difficulty of lithium ions penetrating the SEI film will increase, which is equivalent to an increase in the internal resistance of the battery. The second factor is current rate, high-rate discharge will generate more heat inside the battery and accelerate battery aging ,finally is the cut-off voltage, Overcharge and over discharge of the battery will affect the health of the battery, and inappropriate upper and lower voltage limits will affect the battery. The lower the discharge cut-off voltage, the greater the internal resistance of the battery, resulting in internal heating of the battery, increased side reactions, reduction of battery active materials and collapse of the negative graphite sheet, accelerated aging and capacity decay of the battery. Excessive charge cut-off voltage causes the internal resistance of the battery to increase, and the internal heat of the battery increases. Overcharging causes the negative electrode to produce "lithium precipitation" and the corresponding side reactions increase, which affects the capacity and aging of the battery.

battery SoH characterizes the ability of the current battery to store electrical energy relative to the new battery., which has several performance indicators will be present in the next paragraph.

1. Capacity definition SoH

There are most literatures using battery capacity decay to define SoH. The given SOH definition is as follows:

$$SOH=\frac{C_{\text{aged}}}{C_{\text{rated}}}\times 100\%$$

C_{aged} is the current capacity of the battery; C_{rated} is the rated capacity of the battery, this article will also take this approach, as it is the clearest and most comprehensible way to express it

2. Internal resistance defines SOH

$$SOH=\frac{R_{\text{EOL}}-R_{\text{C}}}{R_{\text{EOL}}-R_{\text{new}}}\times 100\%$$

In the formula: Reol corresponding the internal resistance at the end of the battery life; Rc is the internal resistance of the current battery; Rnew is the internal resistance of the new battery.

The increase in the internal resistance of the battery is an important manifestation of battery aging, and it is also the reason for the further aging of the battery.

3. The number of remaining cycles defines SOH

$$SOH=\frac{Cnt_{\text{remain}}}{Cnt_{\text{total}}}\times 100\%$$

In the formula: Cntremain is the number of remaining cycles of the battery; Cnttota is the total number of cycles of the battery, this kind of representation are usually provided by the battery manufacturer.

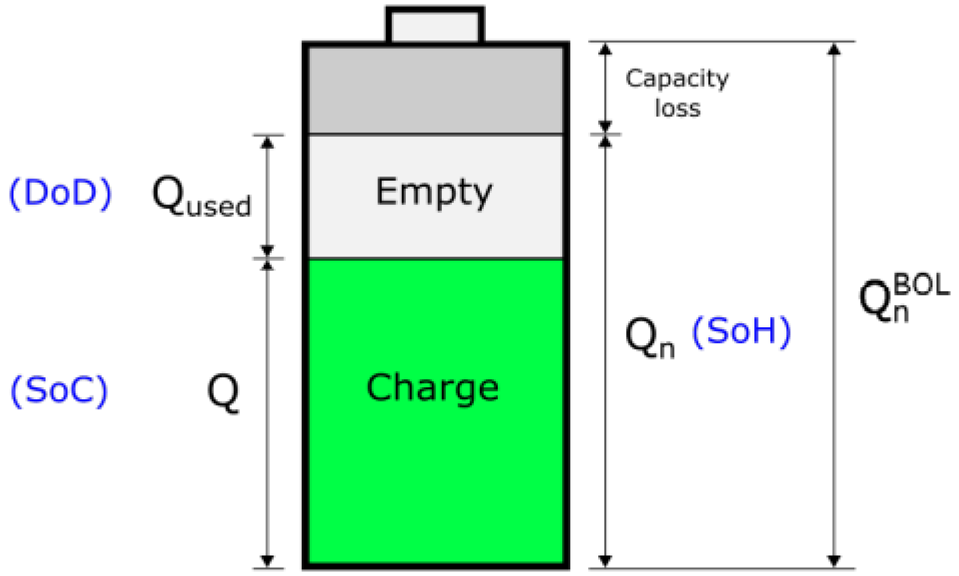


Figure 2.5.2: a summary representation about the three main status parameters (DoD, SoC and SoH) of the battery related to the capacity

Find accurate SoC and SoH is crucial in practical applications:

1. Safety: In order to avoid serious consequences of over-charged and over-discharged, an accurate SOC is very important

2. Reliability: in the case of electric vehicle, driver must be able to estimate the remaining available driving range and be able to choose the driving style from the SoC information.

3. Sustainability: knowing the aging behavior of the battery is very important for environmental protection. Usually, the decomposition of the battery takes a long time. The battery can be recycled or scrapped according to the SOH.

Chapter 3 Available approaches

In the previous chapter, we introduced the hierarchical composition of the battery pack used and the internal chemical reactions. Through the definition of two important indicators SOC and SOH, the approaches we will study and involved in this chapter are introduced.
There exist many kinds of approaches that can be found in the literature to deal with SoC and SoH estimation problem and they can be grouped into the following two macro categories:

• First principles approaches: The estimator is developed according to the physical and or chemical principles. First principles thinking is about acquiring knowledge about a thing by knowing its first causes. First principles thinking is about decomposing a problem or a thing into its most basic elements which can be termed as first causes. The final cause is about the purpose that the things serve.
. • Data-driven approaches: The estimator is a model which maps a given set of data to the target parameter without considering any physical and or chemical principle. It makes strategic decisions based on data analysis and interpretation. A data-driven approach enables companies to examine and organize their data to better serve their customers and consumers.

All methods of these method families rely on the use of some datasets, we used datasets with different schemes and specifications for validation and testing, therefore possible data sources are described in the next paragraph

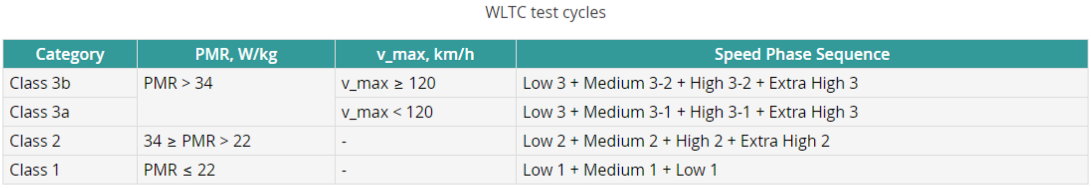
3.1 Data sources

In the automotive field, the leading data source is obtained by considering one or many dynamometer tests. In the circumstance of these tests, typically, the battery is monitored by acquiring measurement of current, voltage and temperature while the target vehicle perform a Dynamometer Drive Schedules (driving cycle). A driving cycle is a collection of speed references with respect to time that capture an average driving behavior of a typical driver in different road conditions and different contexts (i.e. urban, highway, and so on). These tests are important because the resulting data can capture the peculiarity of the application scenario for which the estimator is designed. There exist a lot of driving cycles and every country or company have their own of reference. In European Union the leading one is the Worldwide Harmonized Light Vehicles Test Cycle (WLTC) which has substituted the old New European Driving Cycle (NEDC).[1]

Because of the above reason, In this paper, we use the WLTC loop for the test set under many default criteria.

WLTC, or "Worldwide Harmonized Light Vehicle Test Cycle", is a chassis dynamometer test used to determine light vehicle emissions and fuel consumption and is an important part of the WLTP test procedure. WLTC comprehensive fuel consumption refers to the fuel consumption value of pure fuel or hybrid light-duty vehicles under the WLTP test procedure under the WLTC cycle comprehensive operating conditions, which is convenient for consumers to objectively understand and compare the fuel economy level of different vehicles. The standard is accepted by China, Japan, the United States and the European Union, From July 1, 2021, the test conditions of traditional energy passenger vehicles and plug-in hybrid electric passenger vehicles in China will be switched from NEDC to WLTC. The pure electric passenger car uses the CLTC-P test condition.

There exist many WLTCs, and they are applicable to some specific category of vehicles which are distinguished from their power-to-mass (PMR) ratio. With the highest power-to-mass ratio, Class 3 is representative of vehicles driven in Europe and Japan. Class 3 vehicles are divided into 2 subclasses according to their maximum speed: Class 3a with v_max < 120 km/h and Class 3b with v_max ⩾ 120 km/h.



WLTC test cycles

Category	PMR, W/kg	v_max, km/h	Speed Phase Sequence
Class 3b	PMR > 34	v_max ≥ 120	Low 3 + Medium 3-2 + High 3-2 + Extra High 3
Class 3a		v_max < 120	Low 3 + Medium 3-1 + High 3-1 + Extra High 3
Class 2	34 ≥ PMR > 22	-	Low 2 + Medium 2 + High 2 + Extra High 2
Class 1	PMR ≤ 22	-	Low 1 + Medium 1 + Low 1

Figure3.1.1 WLTC test cycles categories[7]

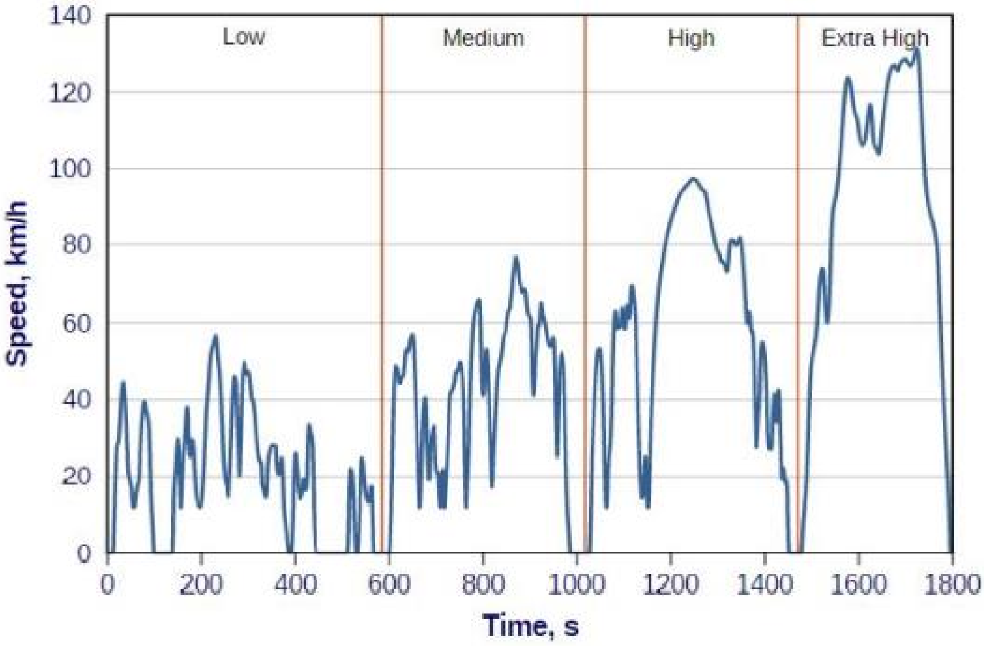


Figure 3.1.2: WLTC cycle for Class 3b vehicles. [7]

In the test of the final application scenario, we also used the following data sets for comparison and verification:

1. NEDC: used as reference cycle for homologating vehicles until Euro6 norm in Europe and some other countries.
2. FTP-75: The FTP-75 (Federal Test Procedure) has been used for emission certification and fuel economy testing of light-duty vehicles in the United States
3. IM240: The IM240 test is a chassis dynamometer schedule developed and recommended by the US EPA for emission testing of in-use light duty vehicles in inspection & maintenance (I&M) programs implemented in several states.

Other types of data can be retrieved by performing some tests directly on the battery system. These tests are typically adopted with the purpose of characterize the batteries properties and making possible to build batteries datasheets. Along with the tests (CC, CC-CV) already described in the previous chapter, there exist the following tests:

1. Pulse Discharge/Charge Test (PDT, PCT), which is also widely used to assess battery characteristics by observing voltage change in pulse duration. Moreover, pulse discharge can be used to simulate the loading of DC power device.

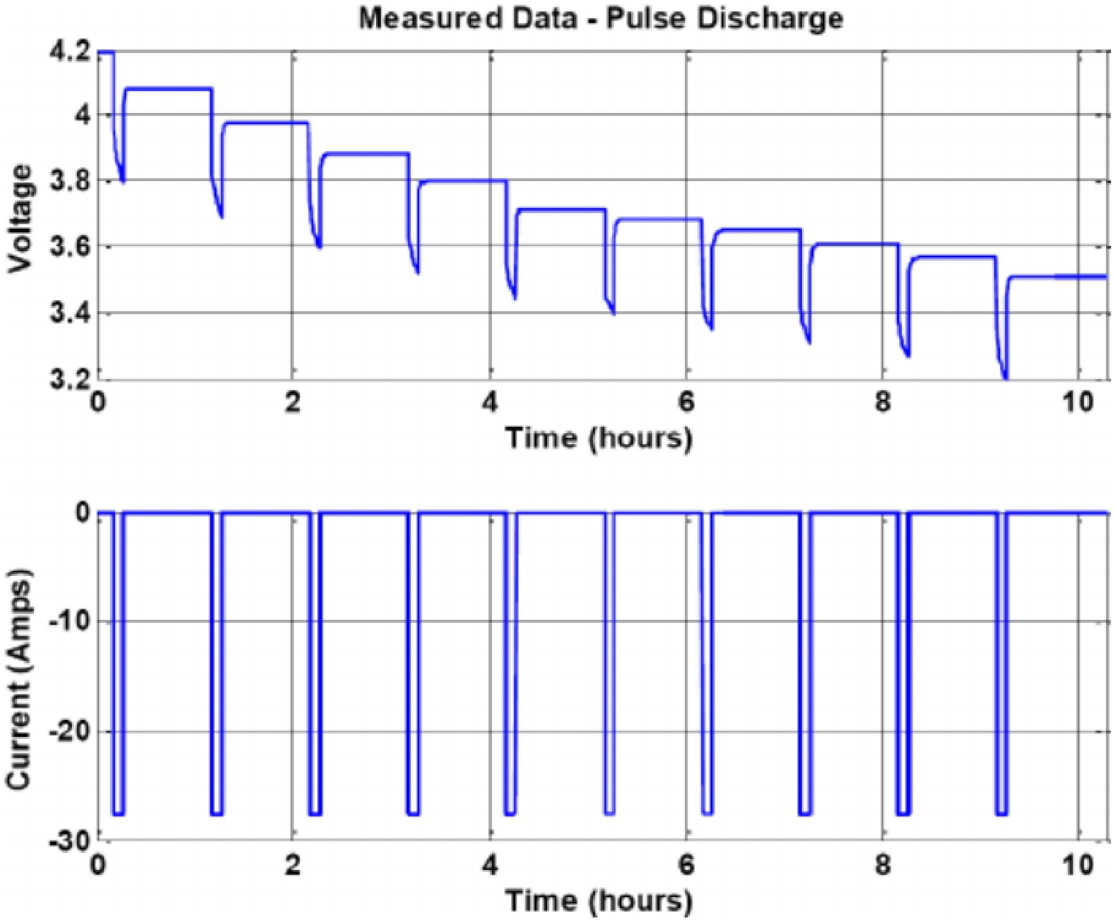


Figure 3.1.3. Example of pulse discharge in 10% increments of SOC [1]

2. Continuous Discharge/Charge Test (CDT, CCT)

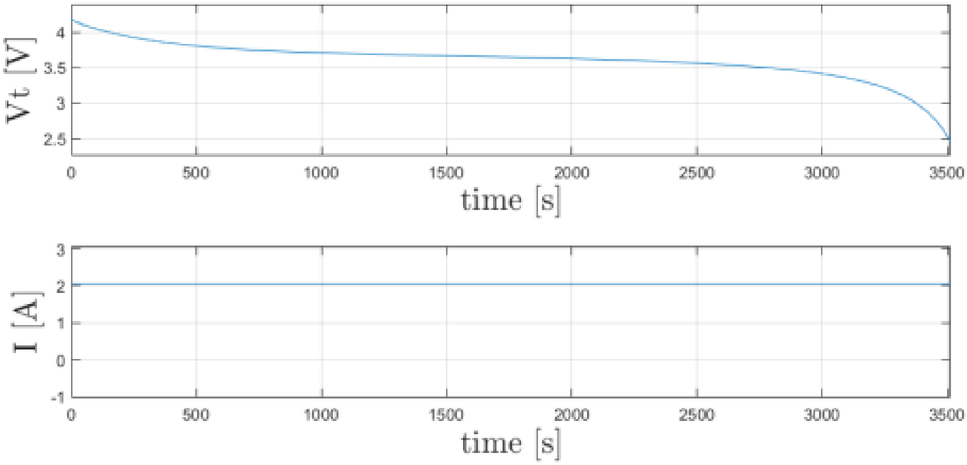


Figure 3.1.4: Example of CDT showing current stimuli and terminal voltage response, executed via simulation on a battery model of a Panasonic

CGR18650AF. Starting from a full charged battery, a 1C continuous discharge current stimulate the battery until it reaches its fully discharged state. [8]

3. Hybrid Pulse Power Characterization test (HPPC).

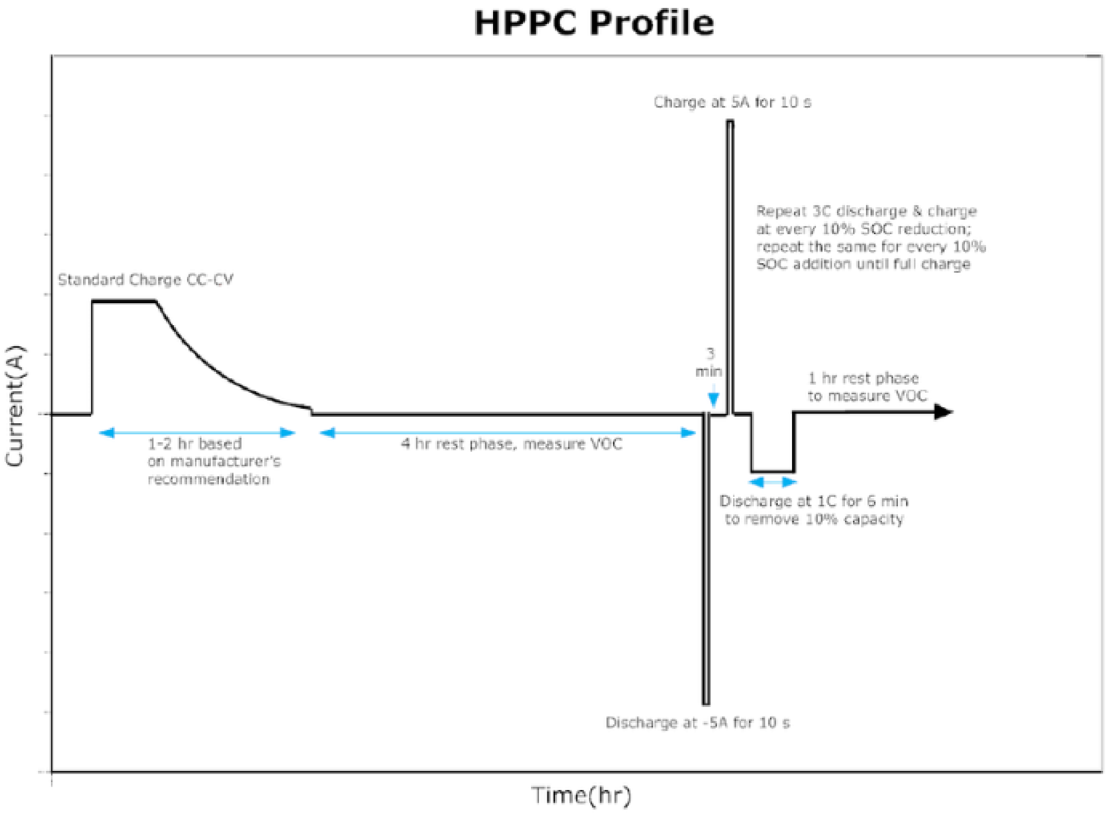

HPPC Profile
Charge at 5A for 10 s
Repeat 3C discharge & charge at every 10% SOC reduction; repeat the same for every 10% SOC addition until full charge
Standard Charge CC-CV
3 min
1 hr rest phase to measure VOC
1-2 hr based on manufacturer's recommendation
4 hr rest phase, measure VOC
Discharge at 1C for 6 min to remove 10% capacity
Discharge at -5A for 10 s
Current(A)
Time(hr)

Figure3.1.5. Hybrid Pulse Power Characterization (HPPC) test profile. [9]

Their execution can be performed using a cycler machine that repeats systematically the tests by alternating full charge and discharge of the battery. In this way is possible to acquire data during the entire life of the battery and make possible to track its age by measuring the capacity at a given iteration (cycle).

Data can be based on real measurements, obtained by testing real systems, or synthesized based on virtual models, by simulations that take into account accurate system models. Since direct physical experiments can be expensive, data based on real data can be found on the web in the form of datasets, which are published by authorities such as universities. Using it to verify the model itself, the model after a certain training can get better accuracy and accuracy.

3.2 First principles approaches

3.2.1 Coulomb Counter

In the coulomb counting method (also known as the Ampere-hour method), the battery SOC is calculated by accumulating the amount of electricity flowing into or out of the battery. The calculation accuracy of the coulomb counting method depends on the initial SOC estimation accuracy and the battery current measurement accuracy (current sensor accuracy). Lithium-ion batteries often have relatively high charge-discharge efficiency and require long-term monitoring, so the coulomb counting method is very convenient for SOC estimation of lithium-ion batteries. Although this method is not suitable for real-time SOC estimation, it can be used to verify the accuracy of results obtained with other estimation methods.

$$SOC(t)=SOC_0-(\int idt)/Qn$$

Among them, SOCo represents the initial SOC value, the current i(t) of the battery pack depends on the state of charge and discharge, and Qn is the rated capacity. The initialized SOC value (SOCo) can be obtained by the OCV method. Although this method is simple and inexpensive, the disadvantages of this method are listed below:

1. Because the coulomb counter is an open-loop estimator, the error of the current sensor will be accumulated by the estimator all the time. As the SOC estimator runs for a long period of time, the accumulated error becomes larger and larger. At the same time, when the current sensor has a relatively large error, it may generate erroneous results faster;
2. When the battery is aging in real time, the capacity of the battery will change, but the coulomb counter cannot detect this problem and will not take other measures for it. As a result, if the actual mode estimated by the battery deviates from the expected mode, the corresponding SOC estimate is also inaccurate;
3. The initial SOC should be obtained by measuring the terminal voltage of the battery pack. Therefore, any error included in the initial estimation method will be carried through the process, and the method has no way to detect or fix the initial error.

For the above reasons, we introduce the modified coulomb counting method.
The coulomb counting method can improve its accuracy by considering the

$$SoC(t) = SoC(t_0) + \int_{t_0}^{t} \frac{\eta_i \cdot I(\tau)}{Q_n} d\tau$$

coulombic efficiency at different temperatures and charging rates. Can be defined as the number of charges output during discharge and the number of charges input during charging or the ratio of discharge capacity to charge capacity.

Qn[A·h] is the nominal battery capacity while ηi is the coulomb efficiency which represents the ratio of the discharge capacity over the charge capacity during a discharge/charge cycle. The coulomb efficiency has to be computed empirically but for simplicity, in the literature can be found that ηi = 1 if "I" is the discharge current, while ηi = η ⩽ 1 if "I" is the charge current. In the expression above battery capacity is stationary over time, and then aging effects are not taken into account. Since the measured signals are sampled over time the function above assume its discretized form:

$$\begin{cases} SoC(k+1) = SoC(k) + \dfrac{\eta_i \cdot \Delta t}{Q_n} \cdot I(k), & k \geq 1 \\ \qquad SoC(k) = SoC_0, & k = 0 \end{cases}$$

Where Qn and "I" are respectively the nominal battery capacity and the charge/discharge current considered in the temporal interval [kΔt,(k + 1) Δt] and Δt is the sampling time interval. Since it is possible to measure the current that is pumped into or requested from the battery, this method results to be simple to implement and can be used both offline and online, but in order to have high accuracy on the resulting value of the SoC it is necessary to have clear knowledge on the measured current. A possible solution to reduce the error is to use this technique together with OCV-based methods that periodically correct the value of the SoC by looking at the existing relation between Open Circuit Voltage and SoC. This correcting technique can be used only in certain conditions that are explained in previous paragraph. Moreover, a crucial aspect regards the SoC initial condition (SoC0) which has to be estimated correctly, otherwise a significant error offset source in the overall estimation is introduced. To avoid issues with the initial SoC, typically the starting scenario refers to a full charged or discharged state of the battery.

The Coulomb Counting approach can be implemented also to estimate the SoH. As discussed before, the process is divided into the following two steps:

$$\begin{cases} Q_{discharged} = \int_0^T I(t)dt \\ SoH = \frac{Q_{discharged}}{Q_{rated}} \times 100\% \end{cases}$$

Starting from a full charged state, the battery is completely discharged, and the relative capacity is obtained through the integration of the discharging current "I". Then, the SoH is computed by dividing the quantity obtained in the previous step by the rated capacity. Also, in this case the accuracy depends on the goodness of the measures of current as explained for the SoC case.

Coulomb counter, hence, suffers in correspondence of a poor-quality sensors (which have to present high resolution and low SNR) and needs helps from OCVSoC method for correction purposes. Moreover, is not always possible to completely charge/discharge the battery online, so, for these reasons, Coulomb Counter is suggested to be used as offline method in the context of a laboratory environment in order to retrieve ground-truth data for other methods

3.2.2 Equivalent Circuit Model based methods

The equivalent circuit model uses a circuit component to form a specific circuit network to characterize the operational characteristics of the circuit. This model establishes the relationship between the external characteristics exhibited by the battery of operation and the internal state of the battery itself.
In order to retrieve SoC and or SoH information, a possibility is to build an estimator based on a battery model which is described by the following general relationship:

$$V_t = f(SOC, I, T)$$

The output voltage is function of SOC, the current I and the battery temperature.

Due to different use cases and application purposes, the function f will vary in complexity according to the difficulty of the application scenario. In particular, when considering a fraction of the capacity rate to draw or deliver current, using a simple model is adequate due to the low dynamic curve, on the other hand when considering high currents (multiples of the capacity rate) or non-constant Current profiles, more complex models are required.

The most complete representation that can be defined in order to describe the behavior of the battery system is an electrochemical model that consider very detailed phenomena happening on microscopic scale in terms of chemical reactions between active elements present into the cells. These types of models are computational expensive since are based on a system of several partial differential equation and present a lot of parameters to be identified, so are not appropriate to be elaborated online on BMS but are useful in the cell design process.

In contrast, the Equivalent Circuit Model (ECM), it uses traditional circuit elements such as resistors, capacitors, and voltage sources to form a circuit network to describe the external characteristics of the power battery. The model uses a voltage source to represent the thermodynamic equilibrium potential of the power battery, and uses an RC network to describe the dynamic characteristics of the power battery. The equivalent circuit model has good applicability to various working states of the power battery, and the state equation of the model can be deduced, which is convenient for analysis and application. Equivalent circuit models have been widely used in new energy vehicle modeling and simulation research and model-based BMS algorithm development. Because it is the simplest solution to consider the macroscopic effects of microscopic phenomena, meanwhile it retain useful information necessary to describe the dynamics of the system. An ECM can be defined at different levels of granularity: starting at the most granular level, it can represent individual cells, individual modules, and at a coarse level, the entire battery pack, respectively. For practical purposes, it is often desirable to use models defined at the battery pack level. In many cases, overall battery pack models are constructed by grouping together cell/module models as basic blocks in corresponding series/parallel fashion. In summary, the model can be viewed as a cell-averaged model for all cells that make up the battery pack.

Among the possible ECMs the most common models are the internal resistance (Rint) ECM, the First and the Second Order Thevenin ECM (FOTM, SOTM), Partnership for a New Generation of Vehicles (PNGV) model and General Non-

Linear (GNL) model. The circuit diagrams of each model, their respective expressions and description are reported in next discussion.

The simplest way to model the battery system is to consider the series between a voltage source representing the Open Circuit Voltage (VOC),.The choice of the voltage source will greatly affect the model accuracy. If an ideal voltage source is selected, the output voltage is equal to OCV and does not change with the current, which is obviously the model with the largest error. Our model adopts an improved model of OCV related to SOC, which is equivalent to the ideal voltage source is replaced by a controlled voltage source with a value equal to OCV(z(t)), reducing the model error, if the temperature T is also considered, OCV(z (t), T(t)) then the error is further reduced. (Look-up table method and interpolation method are easy to handle)

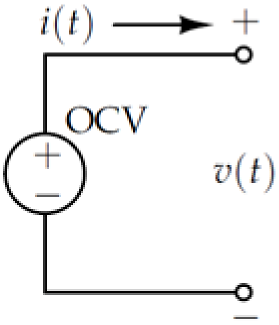


Figure3.2.6. Ideal voltage source
[10]

After adding the internal resistance of the battery, it is natural that the output current has a voltage drop, and the model accuracy is further improved. Note that the resistance of the battery is usually a function of the battery SOC, and it is also a function of the internal temperature of the battery. If these Z and T dependencies are considered in R0, the model accuracy continues to improve.

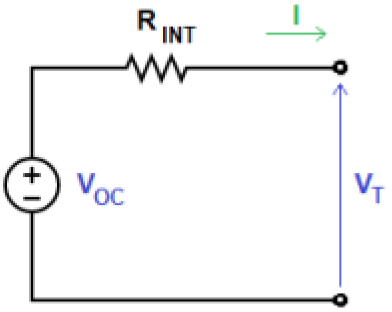
circuit schema	expressions
	$V_T = V_{OC} - R_{INT} \cdot I$

Table 3.2.1: RINT model

According to the thesis work "Electro-Thermal Modelling of Lithium-Ion Battery" conducted by Mohammad Taffel at Politecnico di Torino, this model can be a good choice when the operations performed on the battery stimulate it according to the limit suggested by the cell manufacturer. In these circumstances, it is possible to build the entire battery pack model and starting from the information present in the cell datasheet, construct the OCV-SoC relationship and find the proper internal resistance values, for different temperature and discharge/charge condition. The low complexity of this model allows it to be easily implemented for real time applications.

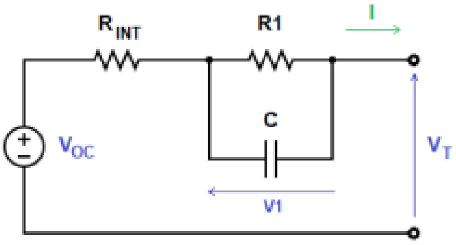
circuit schema	expressions
	$\begin{cases} \frac{dV_1}{dt} = -\frac{V_1}{\tau_1} + \frac{I}{C} \\ V_T = V_{OC} - V_1 - R_{INT} \cdot I \end{cases}$ where $\tau_1 = R_1 \cdot C$

Table 3.2.2: First order Thèvenin model

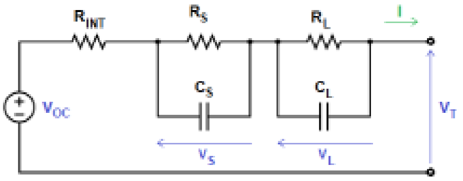
circuit schema	expressions
 R_{INT} R_S R_L I C_S C_L V_{OC} V_T V_S V_L	$\begin{cases} \frac{dV_S}{dt} = -\frac{V_S}{\tau_S} + \frac{I}{C_S} \\ \frac{dV_L}{dt} = -\frac{V_L}{\tau_L} + \frac{I}{C_L} \\ V_T = V_{OC} - V_L - V_S - R_{INT} \cdot I \end{cases}$ where $\begin{cases} \tau_S = R_S \cdot C_S \\ \tau_L = R_L \cdot C_L \end{cases}$

Table 3.2.3: Second order Thèvenin model

The Thèvenin model starts from the Rint schema and introduce respectively in the first (table 3.2.2) and second order (table 3.2.3) versions, one and two RC networks with different time constants. These networks introduce more complexity to the model and permits to describe in more details the short-term and long-term transient behavior of the terminal voltage. Physically RC networks model the diffusion process in the electrolyte and porous electrodes, the charge transfer and double-layer effect in the electrode.

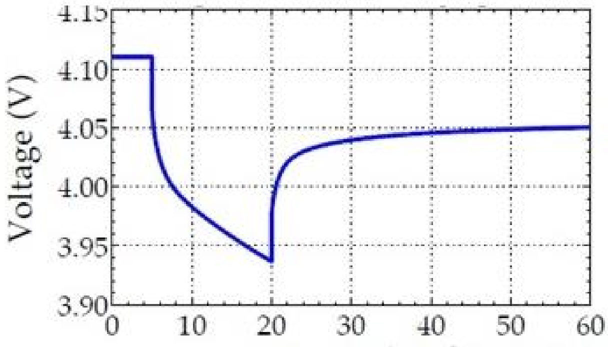


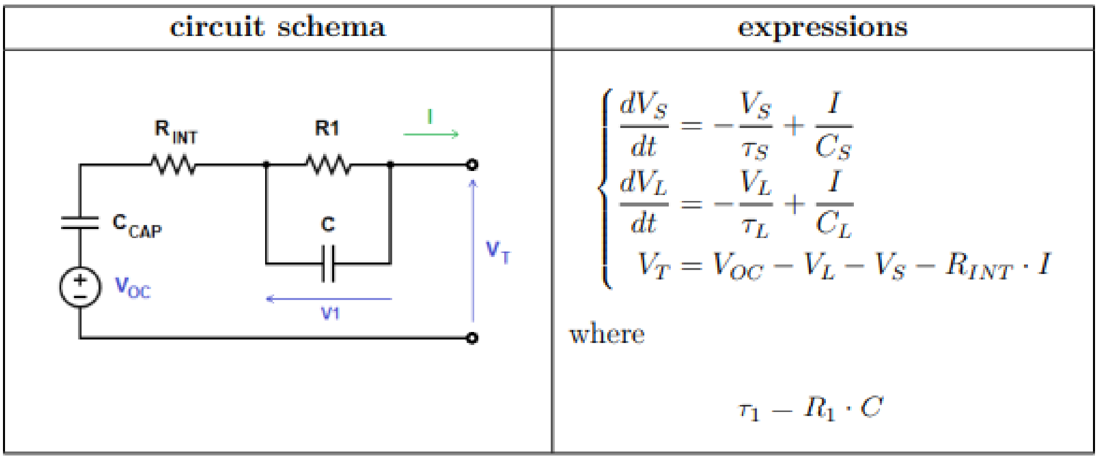
Figure 3.2.7 Polarization effect

[10]

It may be understood with the addition of RC that polarization refers to any deviation of the output voltage of a battery away from OCV due to current passing through the battery. In the previous equivalent circuit model, the instantaneous polarization is modeled by the i(t) × R0 term. But real batteries have more complex behavior, for example, the voltage polarization varies slowly with time when outputting current, and plans to decay slowly with time after stopping the output current. So for the new circuit model it added diffusion behavior.

By adjusting the value of RC, the curve of the voltage over time can be approximated. Likewise, simulation accuracy can be further improved in circuits using two or more RC parallel subcircuits. In general, two RCs can obtain sufficient accuracy without excessive computing power requirements.

If we consider the voltage hysteresis, in actual battery use, we can find that if we discharge the battery to 50% SOC and then stop discharging, the balanced voltage is lower than OCV; and if we charge the battery to 50% SOC and then stop charging, Then the balanced voltage is higher than OCV. These observations suggest a hysteresis in the cell terminal voltage. This results in a range of potentially stable quiescent voltage values for each SOC. (simulation accuracy is seriously challenged)


R INT
R1
I
C CAP
C
V T
V OC
V1

circuit schema	expressions
	$\begin{cases} \frac{dV_S}{dt} = -\frac{V_S}{\tau_S} + \frac{I}{C_S} \\ \frac{dV_L}{dt} = -\frac{V_L}{\tau_L} + \frac{I}{C_L} \\ V_T = V_{OC} - V_L - V_S - R_{INT} \cdot I \end{cases}$ where $\tau_1 - R_1 \cdot C$

Table 3.2.4: PNGV model

PNGV is built starting from the respective Thevenin version by adding the so-called bulk capacitance (CCAP) between the voltage source and internal resistance. According to [1] the presence of an additional capacitance helps to model the variation in term of OCV due to discharge current accumulation effect.

Thanks to this additional component, PNGV model reaches often better result if compared with Rint and first order Thèvenin models. Based on the second-order RC model, the model is connected in series with a capacitor used to simulate the influence of the change of SOC on the open circuit voltage of the battery caused by the time integration of the current.

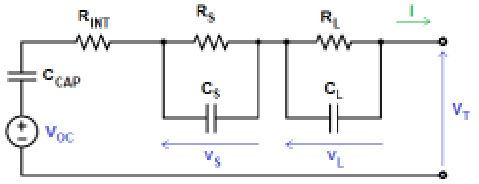
circuit schema	expressions
 RINT R_S R_L I C_CAP C_S C_L V_OC V_S V_L V_T	$\begin{cases} \dfrac{dV_S}{dt} = -\dfrac{V_S}{\tau_S} + \dfrac{I}{C_S} \\ \dfrac{dV_L}{dt} = -\dfrac{V_L}{\tau_L} + \dfrac{I}{C_L} \\ V_{CAP} = \dfrac{1}{C_{CAP}} \int I d\tau \\ V_T = V_{OC} - V_L - V_S - R_{INT} \cdot I \end{cases}$ where $\begin{cases} \tau_S - R_S \cdot C_S \\ \tau_L = R_L \cdot C_L \end{cases}$

Table 3.2.5: GNL model

Finally, GNL model is obtained by adding an extra RC network to PNGV model in order to consider concentration polarization effect. GNL considers the influence of self-discharge factors on the basis of PNGV model, so that the battery characteristics simulated by the model can better reflect the actual properties of the battery.

Every model is characterized by a certain number of parameters that have to be identified. This activity is accomplished a-priori or dynamically, and in both cases empirical data are needed. In the a-priori strategy, parameters are found offline, namely, before the model is deployed. Therefore, at the first, data are acquired during laboratory sessions according to one or more tests described in both paragraphs 2.4 and 3.1. The Open Circuit Voltage is found according to OCV map method (paragraph 3.3.1). During tests all the interesting measures in term of temperature, voltage and current are acquired. SoC and or SoH are monitored and recorded using an offline method, typically Coulomb Counting is applied. The gathered data form a dataset that is spitted in two subsets in order to create an estimation set, used to estimate the parameters, and a validation set, used to validate the final model. A common strategy is to split the whole dataset into 2/3 – 1/3 partitions for estimation set and validation set respectively.

At this point, the model parameters can be found typically by using Least Square method. This estimation technique can be used only if the model relationship is Linear Time Invariant (LTI) and under a certain condition explained later. In these circumstances, and considering all the estimation data, the equation 3.4 can assume the following matrix form expression:

$$y \cong \Phi\theta$$

$$y \in \mathbb{R}^{N_e}, \theta \in \mathbb{R}^{n}, \Phi \in \mathbb{R}^{N_e,n}, N_e \gg n$$

Where y is the output vector containing Ne measurements of the terminal voltage, Φ is the regression matrix containing the measures of SoC, current and temperature and finally θ is a vector containing the n model parameters to be found. The almost equal sign emphasizes the fact that the measures are affected by noise that is omitted in the expression.
The model parameters then are computed by applying the pseudo-inverse of the matrix Φ to the measure vector and it can be performed only if the matrix Φ TΦ is nonsingular

$$\hat{\theta}_{LS} = (\Phi^{T}\Phi)^{-1}\Phi^{T}y$$

Where the matrix ALS = (Φ TΦ) −1Φ T is the pseudoinverse (Φ †) of the regression matrix Φ. Since this method involves the calculation of the inverse of a typically big matrix, LS in this form is used a-priori, namely, offline, because otherwise could be a bottleneck in term of time spent in performing its calculation. This means that the parameters estimated from data are static and they don’t update during the battery operation when the system is working once deployed. Additionally, static parameters don’t allow to adaptively capture any changing phenomena like battery aging effects.
For these reasons, Recursive Least Square (RLS) can be consider for updating the model parameters in an online fashion. Starting from the LS expression is possible to retrieve the following set of equations that are needed to estimate the parameters iteratively:

$$
\begin{cases}
\beta_{k-1} = 1 + \phi(k)^T V(k-1)\phi(k) \\
V(k) = V(k-1) - \beta_{k-1}^{-1} V(k-1)\phi(k)\phi(k)^T V(k-1) \\
K(k) = V(k)\phi(k) \\
\epsilon(k) = y(k) - \phi(k)^T \hat{\theta}_{k-1} \\
\hat{\theta}_k = \hat{\theta}_{k-1} + K(k)\epsilon(k)
\end{cases}
$$

Where V (k0) = αI; k0 is the starting discrete time instant, I is the n×n identity matrix, α ∈ R+ is a parameter to be tuned. If α is almost equal to 1 ˆθk converges rapidly, while if α is much smaller than 1 then ˆθk converges slowly. This technique is quite used and as can be seen it doesn't involve any matrix inverse during the computation process but introduce only an additional cost in term of the number of simple operations. Other possibilities that can be considered in the parameters identification process, which can be found in the literature, consists in utilizing Non Linear Least Square (NLLS), Extended Kalman Filter (EKF), Neural Network (NN), Genetic Algorithm (GA), Fuzzy Logic (FL).

As mentioned at the beginning of this paragraph, the final SoC/SoH estimator is built based on the system model. In the literature the most common adopted model-based estimators are Kalman Filter (KF), Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF). In the following subparagraphs, the theory of each filter is addressed, and some examples of their usage is provided by citing some paper.

3.2.3 Kalman Filter

With the continuous development of sensing technology, robotics, autonomous driving, aerospace and other technologies, the requirements for the accuracy and stability of the control system are getting higher and higher. As a method of optimal state estimation, Kalman filtering is becoming more and more common, such as in the fields of unmanned aerial vehicles and robots.

The Kalman approach, considered the optimal state linear estimator, is based on a probabilistic framework and is associated to a discrete-time LTI system (S), which can be described in term of state space model by means of the following system of equations:

$$S : \begin{cases} x(k+1) = Ax(k) + Bu(k) + w(k) \\ \quad\; y(k) = Cx(k) + v(k) \end{cases}$$

- k ∈ N is a discrete time instant
- x(k) ∈ R n , ∼ (x¯k, Pk) is the state random vector of the system at time instant k with mean value x¯k and covariance matrix Pk
- u(k) ∈ R p is the exogenous input vector of the system at time instant k
- y(k) ∈ R q is the measured output of the system at time instant k
- A ∈ R n,n is the state system matrix
- B ∈ R n,p is the input matrix
- C ∈ R q,n is the output matrix
- w(k) ∈ R n , ∼ N(0, Q) is a sample drawn from a multivariate normal distribution which has zero mean value and Q as covariance matrix; it represents the process noise into the system
- v(k) ∈ R q , ∼ N(0, R) is a sample drawn from a multivariate normal distribution which has zero mean value and R as covariance matrix; it represents the noise afflicting the measurements

For the state estimation algorithm, we can obtain three values of the state quantity: the state predicted value, the optimal estimated value and the real value. The principle of Kalman filtering is to use the Kalman gain to correct the state predicted value and make it approach the real value.

In order to make it easier to understand, the derivation process of the Kalman filter is divided into the derivation of the state estimation covariance, that is, the calculation of the cost function; the second step is the derivation of the Kalman gain matrix and other criteria.
The state prediction value can be obtained from the state prediction equation:

$$\tilde{x}_k^- = A * \tilde{x}_{k-1} + B * u_k$$

The optimal estimate of the state can be obtained from the state update equation:

$$\tilde{x}_k = \tilde{x}_k^- + K(z_k - H * \tilde{x}_k^-)$$

Here zk is the same as y(k),H is the same as C mentioned before.
It can be seen from this equation that the Kalman gain K actually represents the proportion of the model prediction error (Predicted error) and the measurement error (Measurement error) in the process of optimal state estimation, that is, the value range of K is 0 to 1. When K=0, that is, the prediction error is 0, the state value of the system completely depends on the predicted value; and when K=1, that is, the measurement error is 0, the state value of the system completely depends on the measurement value.

Now we define the following equations:

$$e_k^- = x_k - \tilde{x}_k^-$$

$$e_k = x_k - \tilde{x}_k$$

$$P_k^- = E[e_k^- * e_k^{-T}]$$

$$P_k = E[e_k * e_k^T]$$

The first equation represents the prior state error.
The second equation represents the posterior state error.
After that are the covariance between the true value and the predicted value, and the covariance between the true value and the best estimate, respectively.

Next, substitute and derive the equations that have been obtained:

$$\tilde{x}_k = \tilde{x}_k^- + K(H * x_k + v_k - H * \tilde{x}_k^-)$$

$$\tilde{x}_k - x_k = \tilde{x}_k^- - x_k + KH(x_k - \tilde{x}_k^-) + Kv_k$$

$$e_k = (I - KH) * e_k^- - K * v_k$$

It can be known that the estimated error variance matrix is:

$$P_k = E[e_k * e_k^T] = (I - KH) * P_k^- * (I - KH)^T + K * R * K^T$$

The estimation principle of Kalman filter is to minimize the covariance of the optimal state estimation, making it more and more close to the true value.

So we now take the partial derivative of the Kalman gain matrix:

$$\frac{\partial P_k}{\partial K} = -2(P_k^- H^T) + 2K(HP_k^- H^T + R) = 0$$

It can be seen that the Kalman gain matrix under the optimal estimation condition is:

$$K = P_k^- H^T (HP_k^- H^T + R)^{-1}$$

The estimated error variance matrix is:

$$P_k = (I - KH) * P_k^-$$

By simplifying the following equations:

$$e_{k+1}^- = x_{k+1} - \tilde{x}_{k+1}^- = (A * x_k + Bu_k + w_k) - (A * \tilde{x}_k + Bu_k)$$

We get :

$$e_{k+1}^- = A(x_k - \tilde{x}_k) + w_k = Ae_k + w_k$$

From equations mentioned above:

$$P_{k+1}^- = E[e_{k+1}^- * {e_{k+1}^-}^T] = E[(Ae_k + w_k)(Ae_k + w_k)^T]$$

$$P_{k+1}^- = E[(Ae_k)(Ae_k)^T] + E[w_k(w_k)^T]$$

We finally get the prediction covariance matrix:

$$P_{k+1}^- = AP_k A^T + Q$$

Summarized as follows, the state prediction stage:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$$

$$P_k^- = AP_{k-1}A^T + Q$$

State update phase:

$$K_k = P_k^- H^T(HP_k^- H^T + R)^{-1}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$

$$P_k = (I - K_k H)P_k^-$$

[11]
Matrices Q and R can be found according to the a priori statistic information about the process and measurement noises, otherwise have to be found using a trial-and-error approach. Kalman Filter can estimate only SoC: from the perspective of SoH, since it is related to capacity fade or internal resistance degradation, the state vector would contain the capacity or resistance term making the system non-linear.

As a state optimal estimation algorithm, Kalman filter is similar to the observer design thinking in "Modern Control System", which uses the observation quantity combined with the model of the system to estimate the state of the system. However, there are differences between the two. From the previous introduction, it can be seen that the purpose of Kalman filtering is to use Kalman gain to correct the state prediction value, and the purpose of observer design is that when some state information of the system cannot be effectively obtained or difficult to measure, Through state reconstruction, in order to achieve feedback control.
Note that the normal Kalman filter also has its applicability. The limitation of the Kalman filter is that it can only fit a linear Gaussian system. But its biggest advantage is that the amount of computation is small, and the state (and possible measured values) at the previous moment can be used to obtain the optimal estimate of the state at the current moment.

3.2.4 Extended Kalman Filter

The basic Kalman filter algorithm uses a linear random difference equation to estimate the current state, but what if the state estimation relationship and the measurement relationship are nonlinear, and most of the problems in practical use are nonlinear. In addition, normal Kalman states exhibit a unimodal Gaussian distribution. A unimodal Gaussian distribution is still unimodal after being calculated by a linear meter. If it is a nonlinear calculation, the structure is obviously not a unimodal distribution.

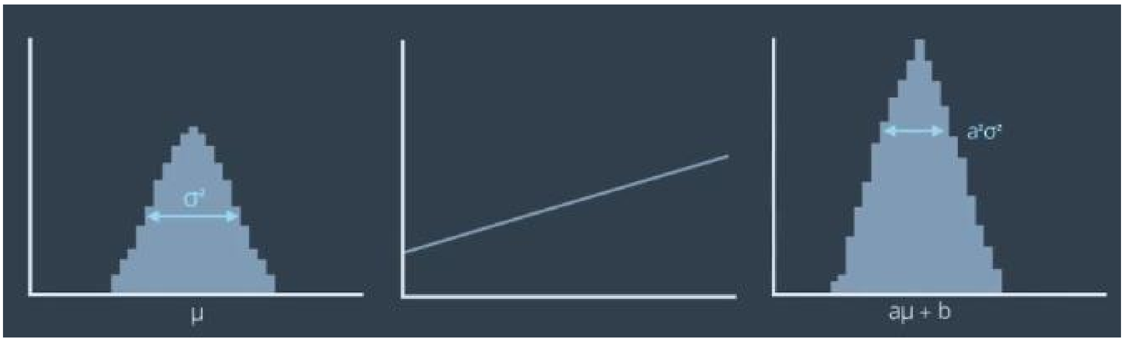


Figure 3.2.8 unimodal distribution

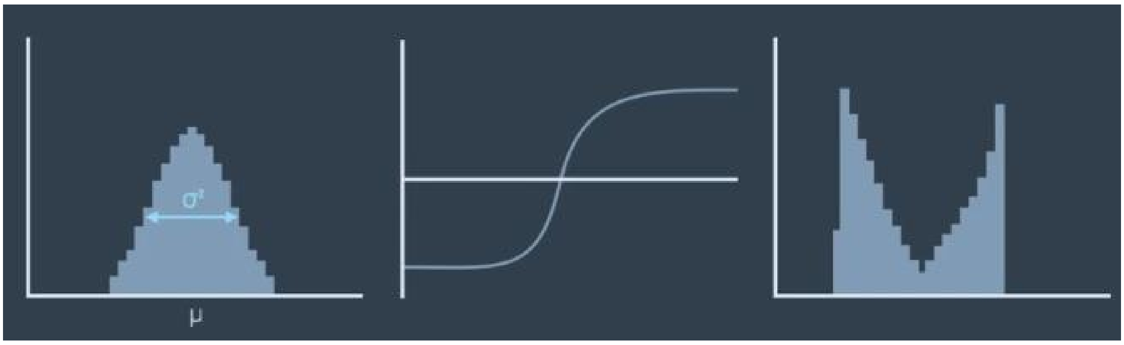


Figure 3.2.9 not a unimodal distribution after non-linear function [12]

So we need to simulate the approximate calculation by extracting the linear algorithm from this nonlinear algorithm. This allows our Kalman filtering algorithm to be used in non-linear environments.

There are many techniques for linearizing nonlinear functions. EKF mainly uses the first-order Taylor Expansion. By obtaining the slope of the nonlinear function at the mean, an approximate linear function is constructed. As shown below:

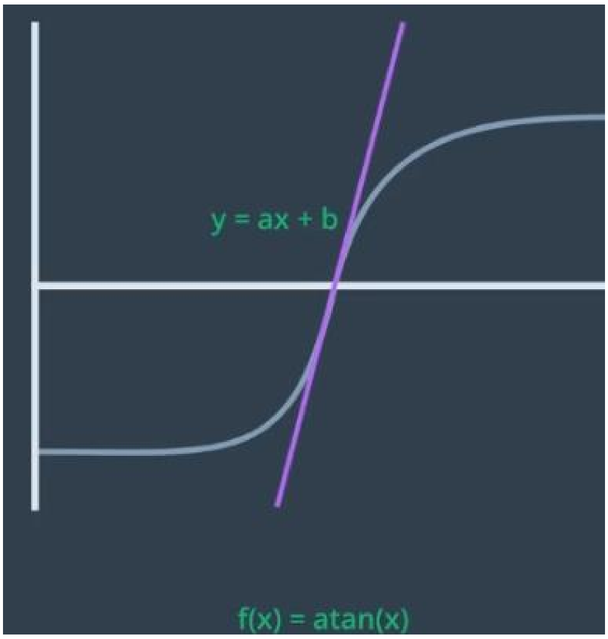


Figure 3.2.10 approximate linear function

These two conditions limit the Kalman filter to a very limited environment so proposed Extend the Kalman filter algorithm to solve nonlinear problems.
So first we need to convert the previous linear system into a nonlinear system：

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + \omega_k \\ z_k &= Hx_k + \nu_k \end{aligned} \quad \Rightarrow \quad \begin{aligned} x_k &= f\left(x_{k-1}, u_k, \omega_k\right) \\ z_k &= h\left(x_k, \nu_k\right) \end{aligned}$$

Where f and h are two well-known differentiable non-linear functions, and the assumptions made on w and v still hold as in the KF setting.
Therefore, we need to linearize the nonlinear motion equations and observation equations with tangents instead. In fact, it is a first-order Taylor expansion at the mean. After a first-order Taylor series expansion, we get the following result：

$$\begin{aligned} A_k &= \frac{\partial f}{\partial x}(\hat{x}_{k-1}, u_k, 0) \\ H_k &= \frac{\partial h}{\partial x}\left(x_k^-, 0\right) \\ W_k &= \frac{\partial f}{\partial \omega}(\hat{x}_{k-1}, u_k, 0) \\ V_k &= \frac{\partial h}{\partial \nu}\left(x_k^-, 0\right) \end{aligned}$$

Among them, Ak and Wk are the Jacobian matrices of the nonlinear system function f after the partial derivatives are obtained at time k, and Hk and Vk are the Jacobian matrices of the nonlinear measurement function h after the partial derivatives are obtained. Then, the extended Kalman filter algorithm can be written as：

1. prediction phase

$$
\begin{aligned}
x_k^- &= f\left(\hat{x}_{k-1}, u_k\right) \\
A_k &= \frac{\partial f}{\partial x}\left(\hat{x}_{k-1}, u_k, 0\right) \\
P_k^- &= A_k P_{k-1} A_k^T + W_k Q_{k-1} W_k^T
\end{aligned}
$$

2. update phase

$$
\begin{aligned}
H_k &= \frac{\partial h}{\partial x}\left(x_k^-, 0\right) \\
K_k &= \frac{P_k^- H_k^T}{H_k P_k^- H_k^T + V_k R V_k^T} \\
\hat{x}_k &= x_k^- + K_k\left(z_k - h\left(x_k^-, 0\right)\right) \\
P_k &= \left(I - K_k H_k\right) P_k^-
\end{aligned}
$$

[13]

The same consideration, as in the KF case, can be done about matrices Q and R. In this context both SoC and SoH can be estimated since the system state vector can contain both SoC and battery capacity or internal resistance.

Using EKF it was possible to identifying along with internal resistance also the nominal capacity, proving the reliability of the data provided by the manufacturer. From the estimated internal resistance signal was possible to notice the remarkable dependence from the SoC. This relationship with SoC was emphasized in the least square approach were, starting from the data it was possible to reconstruct the explicit dependency of the internal resistance from SoC achieving better results with respect to EKF. Once the parameters were found, both EKF and Adaptive EKF (AEKF) were implemented and compared on the SoC estimation task. The difference between the two algorithms lies in the stativity of the covariance matrices of the noises present in the state space model. In the AEKF the Q matrix is adaptively updated in order to follow the estimation error while in EKF is defined offline in function of the matrix R. In both cases matrix R was found by performing a statistical analysis on the measurement error. The updating formula in the AEKF was found by means Maximum Likelihood estimation approach. This work concludes that EKF was able to contain the SoC estimation error into 5% while AEKF allow to reduce the boundary down to 1%.

3.2.5 Unscented Kalman Filter

Lossless Kalman Filter, also known as Unscented Kalman Filter (UKF), is a combination of Unscented Transform (UT) and standard Kalman filter system. The nonlinear system equations are applied to linear assumptions through lossless transformation. standard Kalman system.

UKF uses a statistical linearization technique, which we call unscented transformation. This technique mainly uses linear regression of n points collected in the prior distribution (we call them sigma points). to linearize a non-linear function of a random variable, which is more accurate than Taylor series linearization (a strategy used by EKF) since we are considering the expansion of the random variable. Like the EKF, the UKF is mainly divided into prediction and update.

The basic idea of UKF is Kalman filter and lossless transformation, which can effectively overcome the problems of low estimation accuracy and poor stability of EKF. Because high-order items are not ignored, the calculation accuracy of nonlinear distribution statistics is high. [14]

In the unscented transformation of UKF, we do not need to perform a large number of sampling, but only select a limited number of sampling points, or subject them to nonlinear transformation, and finally the mean and variance of the results after weighted statistical transformation can get a good estimation result.

In summary, the unscented transformation is divided into three steps:

1. Sampling according to certain rules in the original Gaussian distribution
2. The sampling points undergo nonlinear transformation
3. Weighted statistical transformation results

Then the next question is how to sample and assign weights, here is a simple one-dimensional example.

$$X \sim N(\mu, \sigma^2), Y = sin(X)$$

Here is the one-dimensional case, x belongs to a Gaussian distribution, let's estimate y, and in the one-dimensional case, we choose three points, called σ points:

$$x_0 = \mu\text{，}\ x_1 = \mu + \sigma\text{，}\ x_2 = \mu - \sigma$$

Then choose the appropriate weight so that it satisfies:

$$\mu = \sum_i W_i x_i, \sigma^2 = \sum_i W_i (x_i - \mu)^2$$

So we can calculate the mean and variance of y:

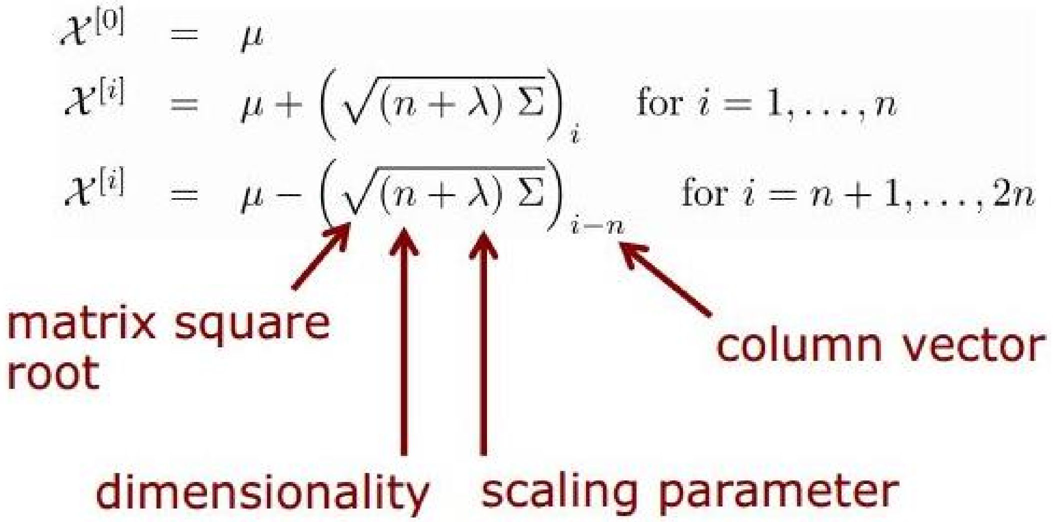
$$\mu_y = \sum_i W_i sin(x_i), \sigma_y^2 = \sum_i W_i (sin(x_i) - \mu)^2$$

Now we can deduce to the more general n dimension, we choose 2n+1 σ points: X(i), and its weight:w(i).

$$\mu = \sum_i w^{[i]} \mathcal{X}^{[i]}$$

$$\Sigma = \sum_i w^{[i]} \left(\mathcal{X}^{[i]} - \mu\right) \left(\mathcal{X}^{[i]} - \mu\right)^T$$

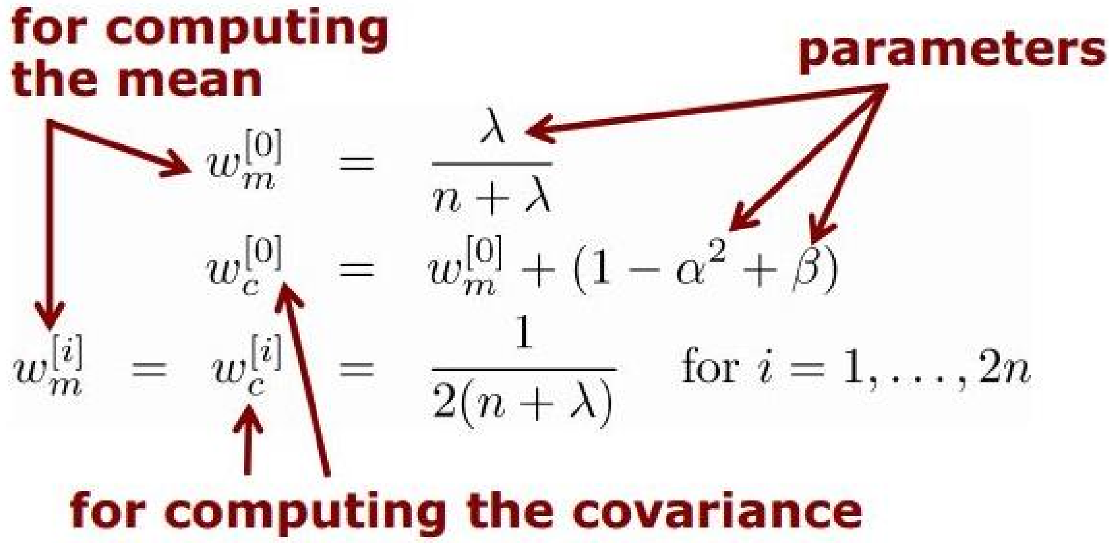
for the sampling point set X(i):



$$\begin{aligned}\mathcal{X}^{[0]} &= \mu \\ \mathcal{X}^{[i]} &= \mu + \left(\sqrt{(n+\lambda)\,\Sigma}\right)_i \qquad \text{for } i = 1, \ldots, n \\ \mathcal{X}^{[i]} &= \mu - \left(\sqrt{(n+\lambda)\,\Sigma}\right)_{i-n} \qquad \text{for } i = n+1, \ldots, 2n\end{aligned}$$

λ represents how far σ is from the mean.

for weight assignment w(i):



$$\begin{aligned} w_m^{[0]} &= \frac{\lambda}{n+\lambda} \\ w_c^{[0]} &= w_m^{[0]} + (1 - \alpha^2 + \beta) \\ w_m^{[i]} = w_c^{[i]} &= \frac{1}{2(n+\lambda)} \qquad \text{for } i = 1, \ldots, 2n \end{aligned}$$

It should be noted that the parameters here satisfy the following conditions:

$$\begin{aligned} &\kappa \geq 0 \\ &\alpha \in (0, 1] \\ &\lambda = \alpha^2(n+\kappa) - n \\ &\beta = 2 \end{aligned}$$

Among them, the more k and α, the farther the sampling point is from the

$$\mu' = \sum_{i=0}^{2n} w_m^{[i]} g\left(\mathcal{X}^{[i]}\right)$$

$$\Sigma' = \sum_{i=0}^{2n} w_c^{[i]} \left(g\left(\mathcal{X}^{[i]}\right) - \mu'\right)\left(g\left(\mathcal{X}^{[i]}\right) - \mu'\right)^T$$

mean.Now we can get the estimation result:

In summary, UKF is also divided into two stages:

1. prediction phase

```
**Unscented_Kalman_filter**$(\mu_{t-1}, \Sigma_{t-1}, u_t, z_t)$:
    $\mathcal{X}_{t-1} = (\mu_{t-1} \quad \mu_{t-1} + \sqrt{(n+\lambda)\Sigma_{t-1}} \quad \mu_{t-1} - \sqrt{(n+\lambda)\Sigma_{t-1}})$
    $\bar{\mathcal{X}}_t^* = g(u_t, \mathcal{X}_{t-1})$
    $\bar{\mu}_t = \sum_{i=0}^{2n} w_m^{[i]} \bar{\mathcal{X}}_t^{*[i]}$
    $\bar{\Sigma}_t = \sum_{i=0}^{2n} w_c^{[i]} (\bar{\mathcal{X}}_t^{*[i]} - \bar{\mu}_t)(\bar{\mathcal{X}}_t^{*[i]} - \bar{\mu}_t)^T + R_t$
```

2. update phase

```
    $\bar{\mathcal{X}}_t = (\bar{\mu}_t \quad \bar{\mu}_t + \sqrt{(n+\lambda)\bar{\Sigma}_t} \quad \bar{\mu}_t - \sqrt{(n+\lambda)\bar{\Sigma}_t})$
    $\bar{\mathcal{Z}}_t = h(\bar{\mathcal{X}}_t)$
    $\hat{z}_t = \sum_{i=0}^{2n} w_m^{[i]} \bar{\mathcal{Z}}_t^{[i]}$
    $S_t = \sum_{i=0}^{2n} w_c^{[i]} (\bar{\mathcal{Z}}_t^{[i]} - \hat{z}_t)(\bar{\mathcal{Z}}_t^{[i]} - \hat{z}_t)^T + Q_t$
    $\bar{\Sigma}_t^{x,z} = \sum_{i=0}^{2n} w_c^{[i]} (\bar{\mathcal{X}}_t^{[i]} - \bar{\mu}_t)(\bar{\mathcal{Z}}_t^{[i]} - \hat{z}_t)^T$
    $K_t = \bar{\Sigma}_t^{x,z} \, S_t^{-1}$
    $\mu_t = \bar{\mu}_t + K_t(z_t - \hat{z}_t)$
    $\Sigma_t = \bar{\Sigma}_t - K_t \, S_t \, K_t^T$
    *return* $\mu_t, \Sigma_t$
```

Similarly to the EKF case, UKF allow to estimate both SoC and SoH.

The comparison between this method and EKF is that its operation speed is slightly slower than EKF, but its accuracy is higher than EKF, and its stability is not as stable as EKF (the higher the matrix dimension, the more unstable), so the application of EKF in industrial production is more extensive, and UKF is mainly used in academia.

[15] [16]

3.3 Data-driven approaches

3.3.1 Open Circuit Voltage map

Open circuit voltage (OCV), that is, the terminal voltage of the battery in the open circuit state, generally has a certain monotonic relationship with the battery SOC. If the data of the open circuit voltage under different SOCs are tested first through experiments, and the relationship curve between the two is fitted, then the SOC of the battery can be estimated according to the measured open circuit voltage.

The estimation accuracy of this method is high, but the disadvantage is also obvious. The open-circuit voltage needs to be measured in the case of an open circuit, and the estimation accuracy can only be guaranteed after standing for a period of time. Therefore, its use range is limited. Generally, before the vehicle starts (the battery has no charge and discharge current), the initial SOC value is estimated by this method.

So we can conclude that it is the simplest method, but also the most accurate one, that can be thought is to construct a one-by-one map (Lookup table) between the Open Circuit Voltage (OCV) and the SoC and or the SoH. Then, by inverting the relationship in correspondence of a specific OCV value, the desired battery state parameter is obtained. The below figure gives an idea of this curve from the perspective of the SoC.

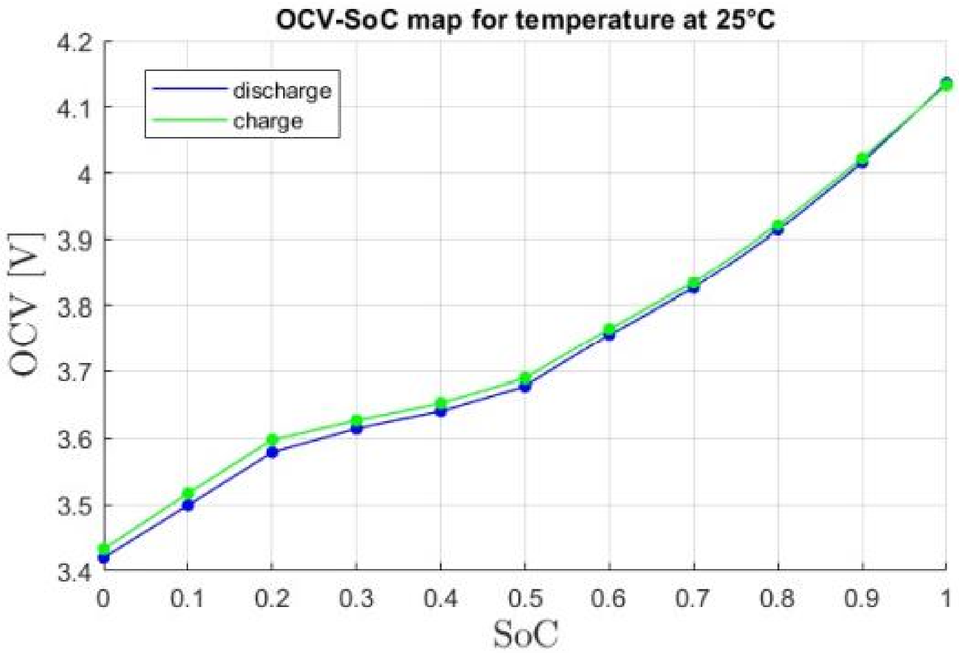


Figure 3.3.1: An example of OCV-SoC graph obtained by representing tabular data retrieved from the Samsung 94Ah prismatic cell datasheet.

As can be seen from the figure above, the existing relationship that exists between OCV and SoC is highly non-linear and depends dramatically on many factors such as temperature, cell aging status (SoH), and current flow direction (charge/discharge), and for this reason the curve, actually, can be thought as a multidimensional surface. Depending on the cell manufacturer, the information regarding the OCV curve can be given in different forms (tabular or graphs) and can be partial (neglecting some dimension like SoH, temperature, or current sign dependency) since the construction of the curve is very time expensive and require high quality hardware to be performed. A typical test adopted to retrieve OCV data points is the PDT (paragraph 3.1). In the context of PDT, by referring to a single cell with a specifying aging status at a given ambient temperature, without loss of generality, the acquisition of a data point happens by reading the terminal voltage once a rest time has passed. This relaxing time is compulsory because it is necessary to wait until electrochemical equilibrium, inside the cell, is reached. In fact, when no current transfer is acting into the cell, after a transient time the measured terminal voltage coincides with the Open Circuit Voltage. Typically, the waiting time needed to get into this situation can be very long (many minutes or hours) and for this reason, OCV measures consist in few points. Often, data refer to a fresh battery (100% SoH) an ambient temperature (25 °C), low C-rate discharge current (charge and discharge behavior is assumed to be equal) and are taken for every 5 or 10% of SoC. The final curve is constructed by means of a curve fitting process which consists in polynomial regression. In the case of SoC the OCV-SoC map can be also given by the following expression:

$$V_{OC} = K_0 + K_1 SOC + \frac{K_2}{SOC} + K_3 \ln(SOC) + K_4 ln(1 - SOC)$$

Where the parameter K0,K1,K2,K3,K4 must be identified on the basis of the acquired data points. Since is expensive to acquire a lot of data in all the operative scenario the accuracy of this method can be penalized. For this reason, OCV method is used as offline method and typically as a complementary or corrective technique. [1]

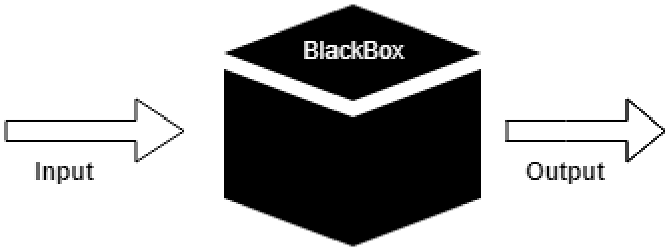
3.3.2 Black box

Black box testing is also known as functional testing, data-driven testing, or functional testing based on requirements specification. This type of testing focuses on testing the functional requirements of the software.

With this test method, the test engineer regards the test object as a black box, completely disregarding the internal logic structure and internal characteristics of the program, and only checks whether the function of the program conforms to its function description according to the "Requirements Specification" of the program. The test engineer does not need to understand the internal structure of the program code, completely simulates the end user of the software product using the software, and checks whether the software product meets the user's needs. The black-box testing method can better and more realistically examine the realization of the functional requirements of the system under test from the user's point of view. In all stages of software testing, such as unit testing, integration testing, system testing, and acceptance testing, black-box testing plays an important role, especially in system testing and validation testing, which cannot be replaced by other testing methods.

Figure 3.3.2 Black box

Since the black box test is based on the user's point of view, the test is carried out from the corresponding relationship between the input data and the output data. Obviously, if there is a problem with the design of the external feature itself or the



specification of the specification is wrong, it will not be detected by the black box testing method.

Lithium-ion battery is a highly complex nonlinear time varying electrochemical system and for this reason is difficult to find a simple model that describe its behaviour in all the operating contexts. Black-Box methods allow to find a model reguardless any physical or chemical principle by simply find a relationship that is able to associate a given input to a given output. In this way is possible to enstablish a direct relationship that relates the input measurements (current, temperature and voltage) to the target signals of interest (SoC and or SoH).
The most common black-box methods employed in the literature are based on machine learning techniques, in particular on Artificial Neural Networks (ANNs). The final estimator is obtained as output of the learning process that is performed on the basis of a given set of data:

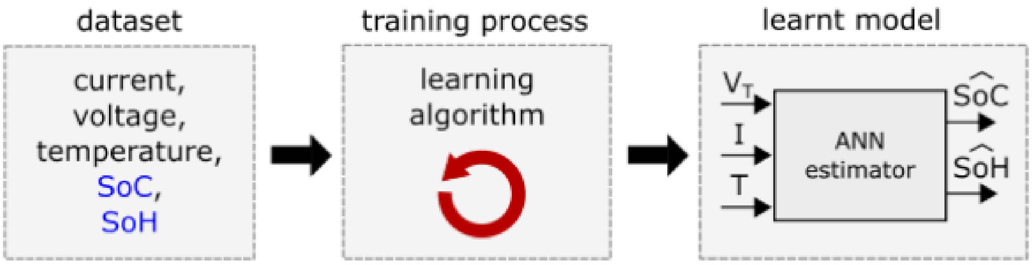


Figure 3.3.3 Basic conceptual schema of ANN estimator development. Blue SoC and SoH are ground truth quantities.

The learning process is performed in a supervised manner, i.e. by providing the target signal that must be estimated along with the input signal samples. By performing experimental tests, the required input signals (current, temperature, terminal voltage) can be obtained, and then using offline methods such as coulomb counters, ground truth SoCs and SoHs can be built. Formally, a dataset can be viewed as a collection of pairs of samples consisting of the input feature vector x and the true value y.
Before training and validation, datasets often need to be preprocessed in order to prepare the data before applying any machine learning algorithms. Common processes in data preprocessing include applying data transformations and data normalization.
We can neither directly use the generalization error as a signal to understand the generalization ability of the model, because the back-and-forth between the deployment environment and the training model is expensive, nor can we use the model's fit to the training dataset as a signal to understand the model's generalization ability. Signals of quantification capability, as the data we obtain are often not clean.
A better way is to split the data into two parts: a training set and a test set. We can use the training set data to train the model, and then use the error on the

test set as the generalization error of the final model in response to real-world scenarios. With the test set, we want to verify the final effect of the model. We only need to calculate the error of the trained model on the test set, and we can consider this error to be an approximation of the generalization error. We only need to let our trained model has the smallest error it could be.

For data and samples of moderate size, we generally use K-fold cross validation method: Divide the initial sampling (sample set X, Y) into K parts, one is reserved as the data for validating the model (test set), and the other K-1 parts are used for training (train set). Cross-validation is repeated K times, one for each validation, and the results are averaged over the K times or using other combinations, resulting in a single estimate.

The advantage of this method is that it repeatedly uses randomly generated subsamples for training and validation, each time the result is validated once, and 10-fold cross-validation is the most commonly used. (Remember that the data used as the validation model is different each time).
In the context of the evaluation of the model, it is required to avoid overfitting by monitoring the behaviour of the generalization error, namely the difference between the training estimation error and the validation estimation error, that has to not increase over time. In fact, a scenario where the training estimation error is low while the validation error is high is suggesting that the model is overfitting the training data, loosing the possibility to perform well on the unseen data provided through the test set.

3.3.3 Long Short Term Memory Recurrent Neural Network

Time recurrent neural network (Recurrent Neural Network) can also be called recurrent neural network. The RNN network adds a "memory" component to the traditional neural network.
In some previous neural network models, it is assumed that the training data are independent of each other, but in many practical applications, the data are interdependent. It is a mutual inheritance relationship. For another example, in a

spatial structure scenario, there is a spatial combination relationship between data, and the overall data can be divided into small local structures, and the properties of the whole can be deduced from the local small structures.

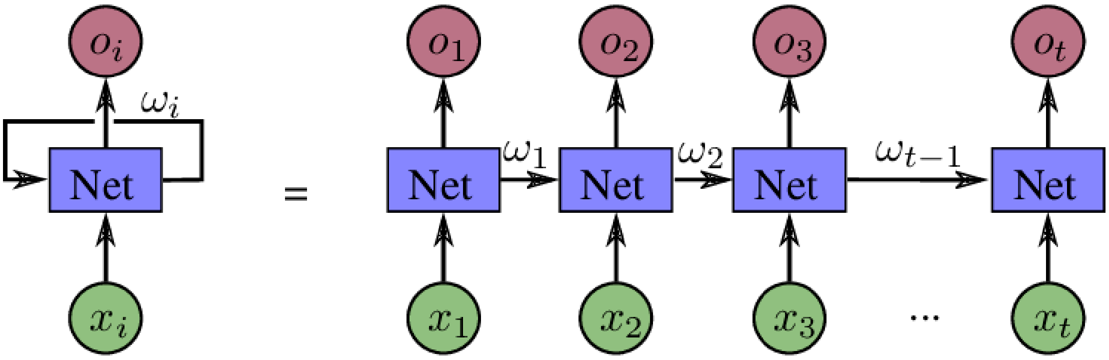


Figure 3.3.4 RNN structure

From the picture above we can see it consists of an input layer, a hidden layer and an output layer. Its characteristic is that a weight matrix W is considered, and the value of the hidden layer of the recurrent neural network depends not only on the current input x, but also on the value of the previous hidden layer. The weight matrix W is the weight of the previous value of the hidden layer as the input of this time.

xi is the input information at time t, oi is the output information at time, we can see that the neuron will recursively call itself and pass the information at time t-1 to time t.

Recurrent Neural Networks work well in many situations, especially when it comes to the analysis of short time series data. However, notice the emphasis on "short" earlier, why is this? The simple recurrent neural network shown in the figure above has a "disadvantage", long-term dependency problem: the recurrent neural network can only handle the situation where we need a closer context. When the information we need to obtain from the above is extremely short, RNN can receive the information from the previous moment and complete the entire information, but it is difficult to complete the information if the required information is in the layers from the first levels. Not only that, compared to the general neural network, this simple RNN is prone to two notorious problems in neural networks: the vanishing gradient problem (the weight/bias gradient of the neural network is extremely small, causing the neural network parameter 's adjustment rate drops sharply) and the gradient explosion problem (the weight/bias gradient of the neural network is extremely large, which causes the neural network parameter adjustment range to be too large and overcorrect).

Therefore, we introduced LSTM. LSTM has been used to solve the long-term dependency problem ubiquitous in general recurrent neural networks since it was designed. Using LSTM can effectively transfer and express information in long-term sequences without causing long-term problems or useful information is ignored (forgotten). At the same time, LSTM can also solve the vanishing/exploding gradient problem in RNN.

1. In a time series, not all information is equally effective, and there are "keywords" or "key frames" in most cases.
2. We will "automatically" summarize the content of the read part when reading from beginning to end and use the previous content to help understand the following.

Based on the above two points, the designers of LSTM proposed the concept of "long short-term memory" - only part of the information needs long-term memory, and some information can not be remembered. At the same time, we also need a mechanism that can dynamically handle the "memory" of the neural network, because some information may be very valuable at the beginning, and then gradually decline in value. At this time, we also need to let the neural network learn to "forget" specific information.

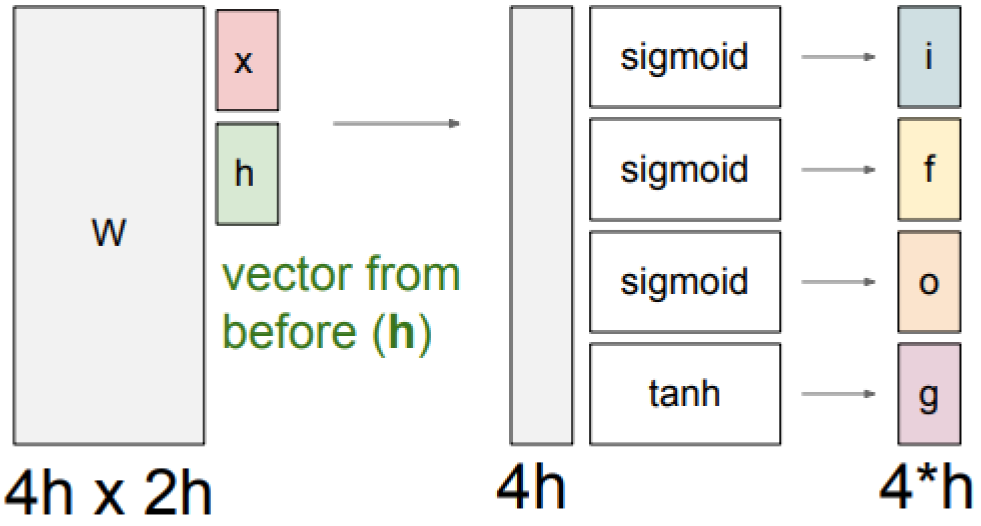


Figure 3.3.5 Long Short Term Momory[17]

From the picture above we can see that there are four ‘gates’ we have to introduce :

1. f: Forget gate, Whether to erase cell
2. i: Input gate, whether to write to cell
3. g: Gate gate (?), How much to write to cell

4. o: Output gate, How much to reveal cell

$$\begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix}$$

$$c_t = f \odot c_{t-1} + i \odot g$$

$$h_t = o \odot \tanh(c_t)$$

Where ct represents the cell state, ht is from the RNN which represents the new state.

LSTM-RNN, or simply LSTM networks are very samilar as RNN but the inner layer, called cell, has a diffrent structure. A new inner state variable is introduced which act as a very memory cell and thanks to the present of inner gates it is given to the network the ability to decide when remember an information as well as when forget it and finally how much of the memorized information to use. All these behaviours happens at every time step and are encoded through the following system of equation:

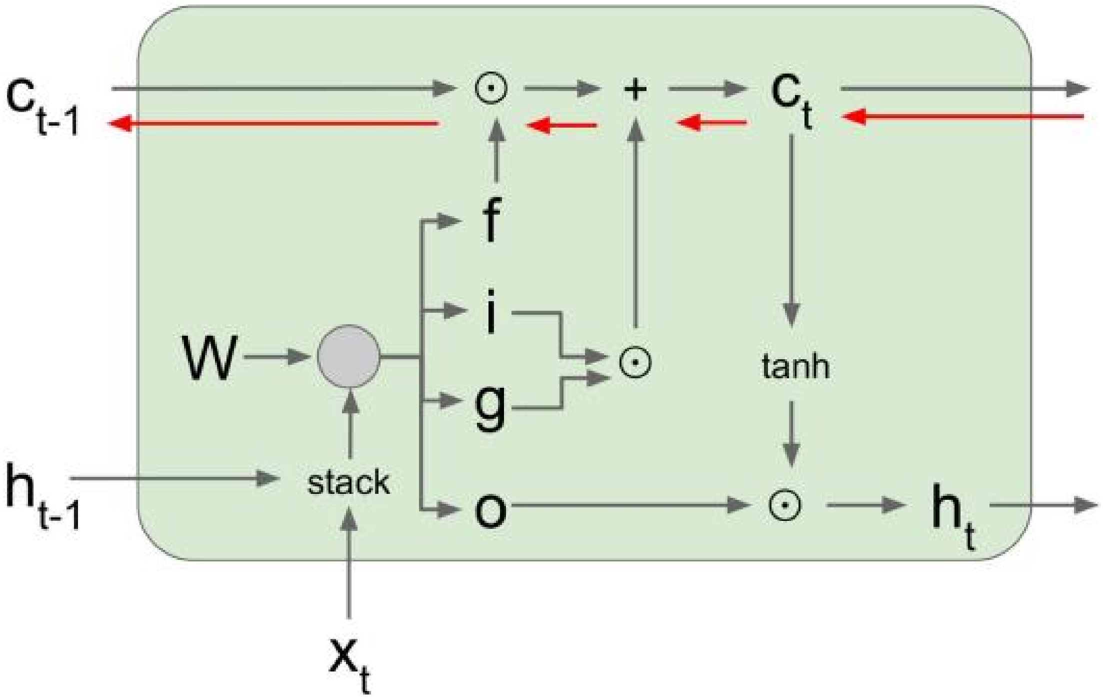


Figure 3.3.6 LSTM cell inner structure.

Be careful that when backprop from ct to ct-1 only elementwise multiplication by f, no matrix multiply by W.

Compared with a simple RNN, a single LSTM unit has a more complex internal structure and input and output. This architecture gives the ability to the network to learn long time dependencies, thanks to the presents of a new inner cell state and the presence of the gates, solving the short-term memory issue of the classic RNNs. Moreover, the presence of the simple path associated with the inner cell state (the straight line that links the previous inner cell state with the current one), makes possible at the same time to simplify the backpropagation step and to dramatically reduce the chance of having problem with vanishing or exploding gradient. There exist many different representations of the LSTM cell that re-arrange the inner gates in a different configuration but the result remains the same. A particular reformulation, which is commonly used, is called Gated Recurrent Unit (GRU) and it uses only two gates called reset and update with the purpose of summarizing the operations that input, forget and output gates do in the classical LSTM cell.

3.4 Summary

From the point of view of usage, methods suitable for the scene need to be adopted according to different application situations, and these estimation methods can be divided into online and online categories. offline technology can be used to acquire data during battery testing in a laboratory environment, where high-quality sensors can be used to ensure the accuracy of data acquisition and can guarantee better control of battery operation. In addition, we can also support other methods by providing ground truth data. The offline method is a Coulomb counter and OCV map. The online method is supported by the offline method in the development step. The combination of offline and online methods can greatly improve the real-time application in practical situations, and can also be implemented on the BMS of EV vehicles. The online methods are KF, EKF, and UKF from the family of first-principles methods, and from the data-driven family RNN and LSTM-RNN.

It is worth noting that the method using artificial intelligence may obtain more accurate prediction results than the Kalman filter, but it is extremely dependent on the sample quality of the training set data collection, which also means that the method will take more time, more energy and even more money in the preparatory work. In contrast, the Kalman filter is a more balanced solution that provides relatively accurate prediction results while reducing the complexity of the model and its dependence on input data.

Chapter 4 Model based design

The previous chapters are aimed to give an overview on the whole battery, especially the SoC and SoH parts of the battery.

At this point, the target of this thesis is presented: design an algorithm that be able to estimate the state of the battery. The application scenario emulates an electric vehicle (in our case Fiat panda after retrofit) that undergoes to a virtual test bench for a given standard driving cycle, for ex: WLTP, NEDC ⋯, at a given ambient temperature, the battery which going to modelling is CE32BNCD-50Ah Li-ion battery, jointly produced by Samsung and BYD, the first step in designing the algorithm is to estimate the parameters of the battery to build the battery model, these steps refer to my colleague Domenico CARLUCCI [1]

The figure below represents the final goal, which is design an estimate algorithm given by a 'black box' form, the available information for the input is the battery terminal voltage, battery input current, battery average temperature, they are the parameters that can be easily detect by apply corresponding sensors.

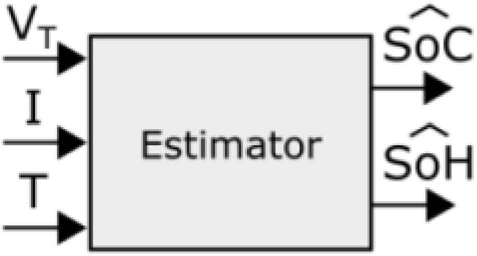


Figure 4.1: Conceptual schema of the estimator.

4.1 virtual test bench

Virtual test bench aimed to provide the behavior of an electric vehicle execute a given velocity profile at certain ambient temperature. In order to reach the goal of this thesis, a simply one-dimensional degree of freedom model has been considered enough to simulate the behavior of the vehicle since the steering system has very little impact on the battery.
the overview of the test bench is shown in the figure below:

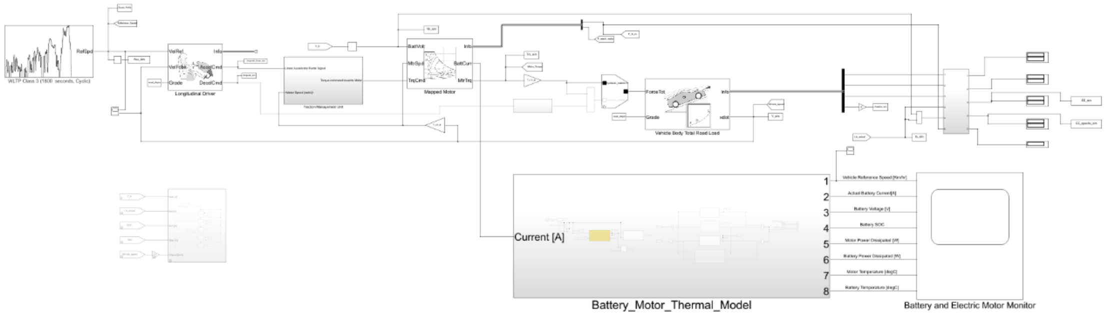

RefSpd
WLTP Class 3 (1800 seconds, Cyclic)
VelRef
VelFdbk
Grade
Info
AccelCmd
DecelCmd
Longitudinal Driver
Traction Management Unit
BattVolt
MtrSpd
TrqCmd
Info
BattCurr
MtrTrq
Mapped Motor
FrcTot
Grade
Info
xdot
Vehicle Body Total Road Load
Current [A]
Battery_Motor_Thermal_Model
1
2
3
4
5
6
7
8
Vehicle Reference Speed [Km/hr]
Actual Battery Current[A]
Battery Voltage [V]
Battery SOC
Motor Power Dissipated [W]
Battery Power Dissipated [W]
Motor Temperature [degC]
Battery Temperature [degC]
Battery and Electric Motor Monitor

Figure 4.1.1: overview of the virtual test bench

Driver behavior can be thought as a controller receive a certain number of inputs from the vehicle and the environment and output the control signal (in this case the position of the acceleration pedal and the braking pedal)

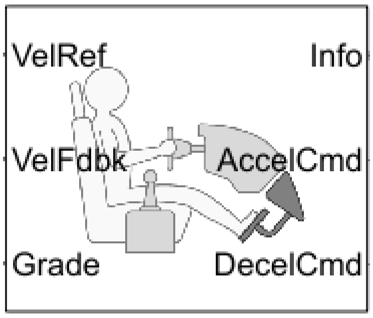


Figure 4.1.2: module that simulates driver behavior

Electric motor, the power system of the electric vehicle is represented by map motor

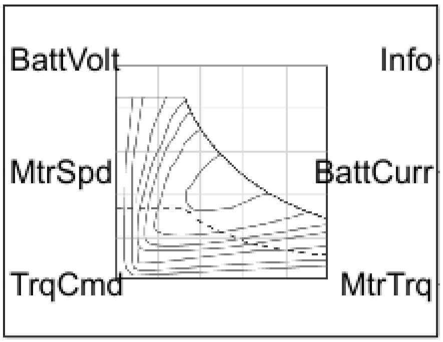


Figure 4.1.3: module simulates electric motor

The longitudinal behavior of the vehicle can be described via a coast-down method, it has detailed describe in [18]

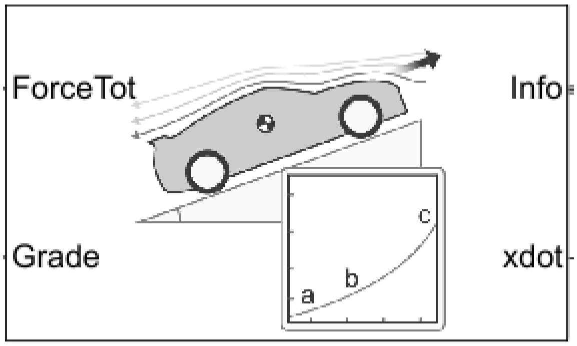


Figure 4.1.4:module simulates longitudinal behavior

The most important part: battery will be described by the data sheet model provided by Simulink, which will be detailed explanation in the next paragraph.

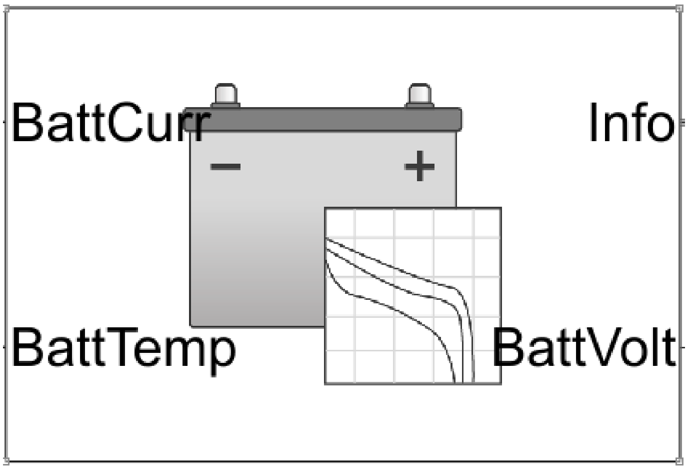
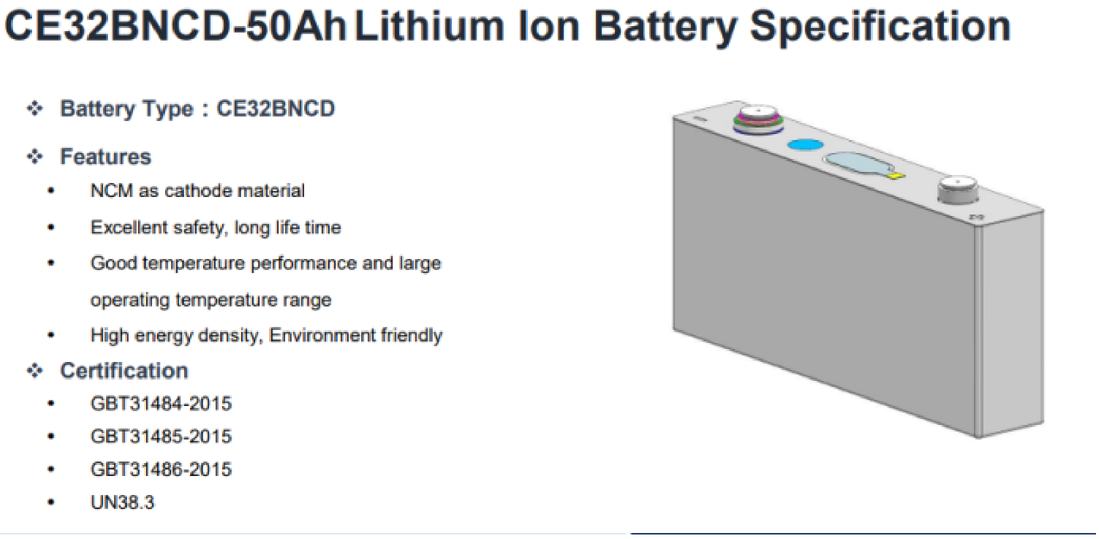


Figure 4.1.5: datasheet battery model

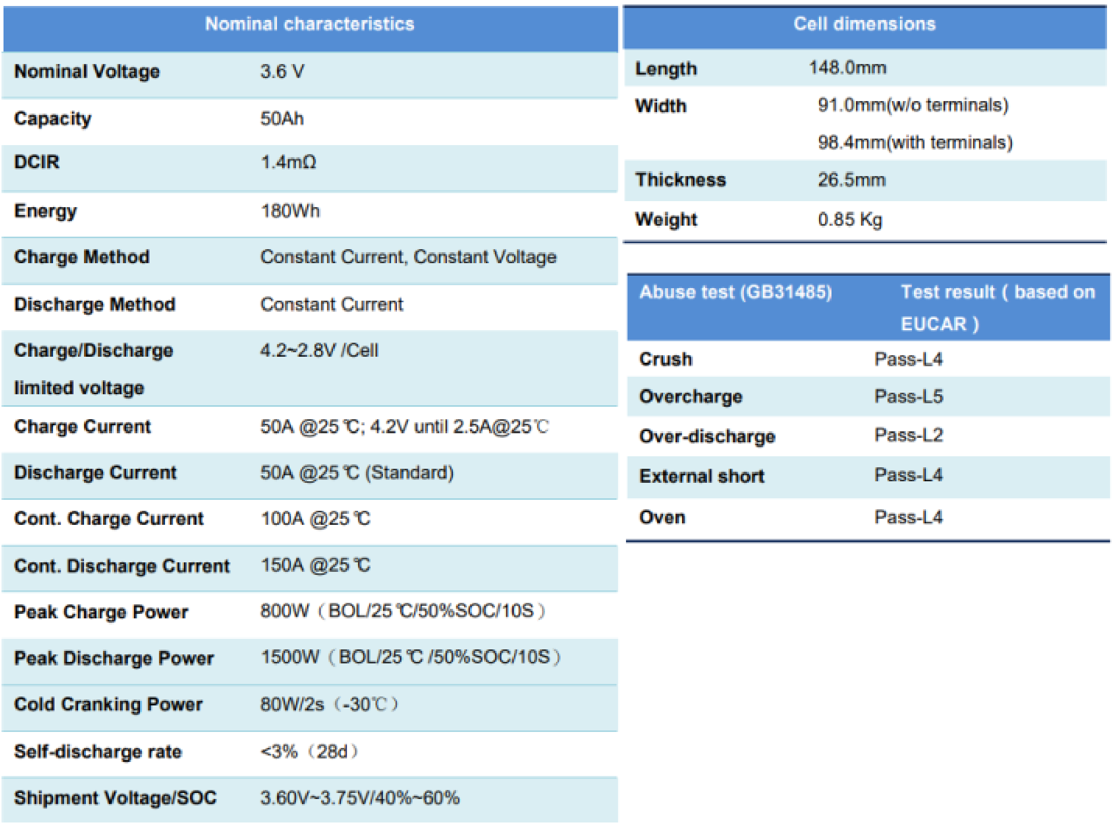
4.2 Battery Model

The battery pack to be modeled is composed by double battery module with 28 prismatic CE32BNCD-50Ah cells combined in a series fashion. The main parameters of the cells are reported below:



CE32BNCD-50Ah Lithium Ion Battery Specification

- Battery Type：CE32BNCD
- Features
  - NCM as cathode material
  - Excellent safety, long life time
  - Good temperature performance and large operating temperature range
  - High energy density, Environment friendly
- Certification
  - GBT31484-2015
  - GBT31485-2015
  - GBT31486-2015
  - UN38.3



Nominal characteristics	
Nominal Voltage	3.6 V
Capacity	50Ah
DCIR	1.4mΩ
Energy	180Wh
Charge Method	Constant Current, Constant Voltage
Discharge Method	Constant Current
Charge/Discharge limited voltage	4.2~2.8V /Cell
Charge Current	50A @25℃; 4.2V until 2.5A@25℃
Discharge Current	50A @25℃ (Standard)
Cont. Charge Current	100A @25℃
Cont. Discharge Current	150A @25℃
Peak Charge Power	800W（BOL/25℃/50%SOC/10S）
Peak Discharge Power	1500W（BOL/25℃ /50%SOC/10S）
Cold Cranking Power	80W/2s（-30℃）
Self-discharge rate	<3%（28d）
Shipment Voltage/SOC	3.60V~3.75V/40%~60%

Cell dimensions	
Length	148.0mm

Width	91.0mm(w/o terminals)
	98.4mm(with terminals)

Thickness	26.5mm
Weight	0.85 Kg

Abuse test (GB31485)	Test result (based on EUCAR)
Crush	Pass-L4
Overcharge	Pass-L5
Over-discharge	Pass-L2
External short	Pass-L4
Oven	Pass-L4

Figure 4.2: main characteristics of the Samsung SDI 94Ah cell.

Battery model is based on a Simulink Specialized Power Systems element block called "Generic Battery Model" (GBM), which is a Rint model described in chapter 3.1, the equivalent electrical model is consist of a voltage source (open circuit voltage) and the internal resistance which value is depending on Soc, Soh and temperature, the states of the battery is calculated by coulomb counting method as described in chapter 3.2. Next figure shown the logic of this block.

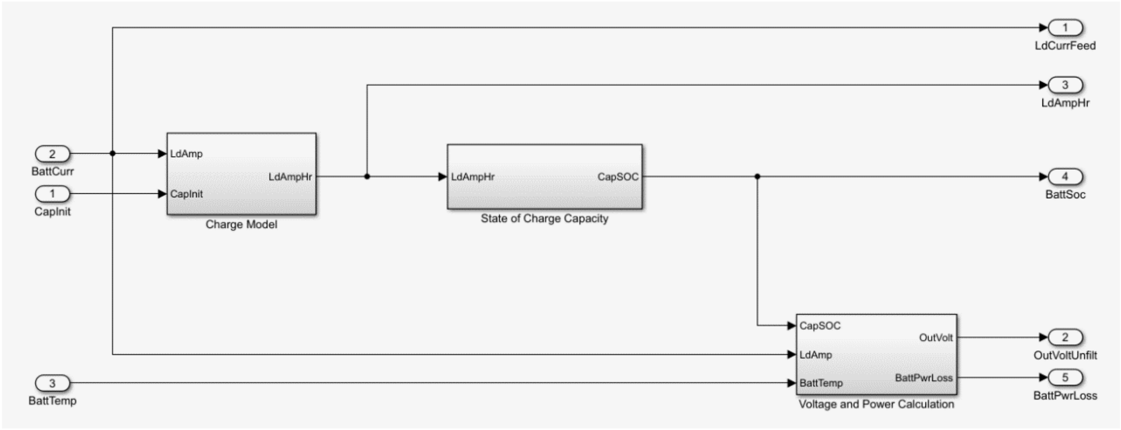


Figure 4.3: algorithm inside the GBM

The main assumptions of the model are the same as [19]

- The internal resistance is assumed to be constant during the charge and discharge cycles and does not vary with the amplitude of the current.

- The parameters of the model are derived from the discharge characteristics. The discharging and charging characteristics are assumed to be the same.

- The capacity of the battery does not change with the amplitude of the current (There is no Peukert effect).

- The self-discharge of the battery is not represented. It can be represented by adding a large resistance in parallel with the battery terminals.
- The battery has no memory effect.

The parameters needed to the GBM can be taken from the cell datasheets, [19] through tabular data and discharge curve graphs at different temperature. The information then can be adapted in order to represent a given battery pack with a given cell topology. In the following table, all the parameters needed to the GBM to model the considered battery pack are reported:

		cell	pack(28S2p)	unity
parameters	nominal voltage	3.68	103.04	V
	rated capacity	50	100	Ah
	initial SOC	100	100	%
	battery response time	855	855	s
discharge	maximum capacity	50	100	Ah
	cut-off voltage	2.8	78.4	V
	fully charged voltage	4.218	118.1	V
	nominal discharge current	16.67	33.33	A
	internal resistance	0.00014	0.00196	Ohm
	capacity at nominal voltage	24.2	48.4	V
	exponential zone[Z,Q]	[3.71 21.5]	[103.88 43]	[V Ah]
Temperature	initial cell temperature	25	25	℃
	nominal ambient temperature	25	25	℃
	second ambient temperature	-25	-25	℃
	maximum capacity	39.9	79.8	Ah
	initial discharge voltage	4.238	118.664	V
	voltage @ 90% capacity	4.08	114.24	V
	exponential zone	[3.82 2.3]	[106.96 4.6]	[V Ah]
	thermal resistance cell-amb	0.06	0.06	℃/W
	thermal time cost cell-amb	1000	1000	s
	heat loss difference	0	0	W
Aging	initial battery age	0	0	cycle
	aging model sampling time	1.80E+04	1.80E+04	s

	ambient temperature Ta1	25	25	℃
	capacity @ EOL	40	80	Ah
	internal resistance @ EOL	0.00028	0.00392	Ohm
	charge current[Ic Icmax]	56	260	A
	discharge current[Id Idmax]	208	720	A
	cycle life @100 DOD	4000	4000	cycle
	cycle life @25 DOD	28000	28000	cycle
	cycle life @100 DOD Idmax	3551	3551	cycle
	cycle life @100 DOD Icmax	3900	3900	cycle
	ambient temperature Ta2	5	5	℃
	cycle life @ 100 DOD Ic and Id	3793	3793	cycle

Table 4.1: Generic battery model parameters. See GBM documentation for parameters details [3].

In order to evaluate the performance of the GBM, using its Simulink implementation, a discharge test is performed, and the results are compared with the ones present in the CNR report according to the same test procedure:

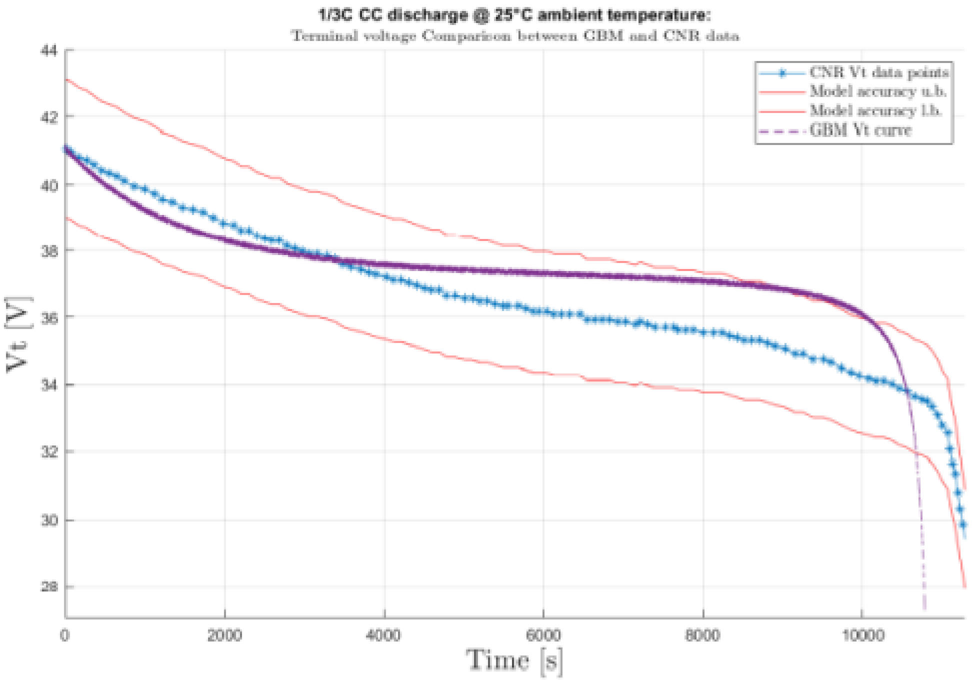


Figure 4.4: The terminal voltage response produced by the GBM is adapted to the 10S1P configuration in order to be comparable with the one provided by the CNR. [26]

The obtained results, according to the characteristics and the stated accuracy of the GBM, are good enough to confirm the validity of the parameters setting.

As mentioned in the previous paragraph (4.1), the purpose of the GBM is to retrieve information that are needed to the datasheet model of the dynamic simulator, to work properly at different aged states. In order to obtain OCV and internal resistance values for different temperature, SoC and SoH, a First Order Thevenin (FOTM) ECM is identified through the acquired data obtained performing some tests on the GBM. The obtained parametric model is also adopted to support the development of the EKF which is addressed in chapter 5.

The tests are performed using the following sets of ambient temperatures and equivalent full cycles (Efc) which represent a specific battery aged status in terms

of the number of complete charge-discharge cycles (0 represent battery BOL and 4000 is the EOL according to datasheet data):

$$T_a \in \{15{,}20{,}25{,}30{,}35{,}40\} \qquad (°C)$$
$$Efc \in \{0{,}400{,}800{,}...{,}4000\} \quad (cycle)$$

The next paragraphs enter the details on how the needed parameters are obtained.

4.2.1 SoH ground truth evaluation

As mentioned in chapter 2.5, the maximum capacity will be decreased cause of the aging behavior, so called capacity fade phenomenon.
In order to get all the SoH at a given temperature and Effective fully discharged cycle, a 1C constant discharge test is applied, the starting point is considered Efc=0 and SoH=100, once the Efc reach 4000 cycles, means the SoH of the battery is lower than 80%, is considered the End of Life (EOF) of the battery

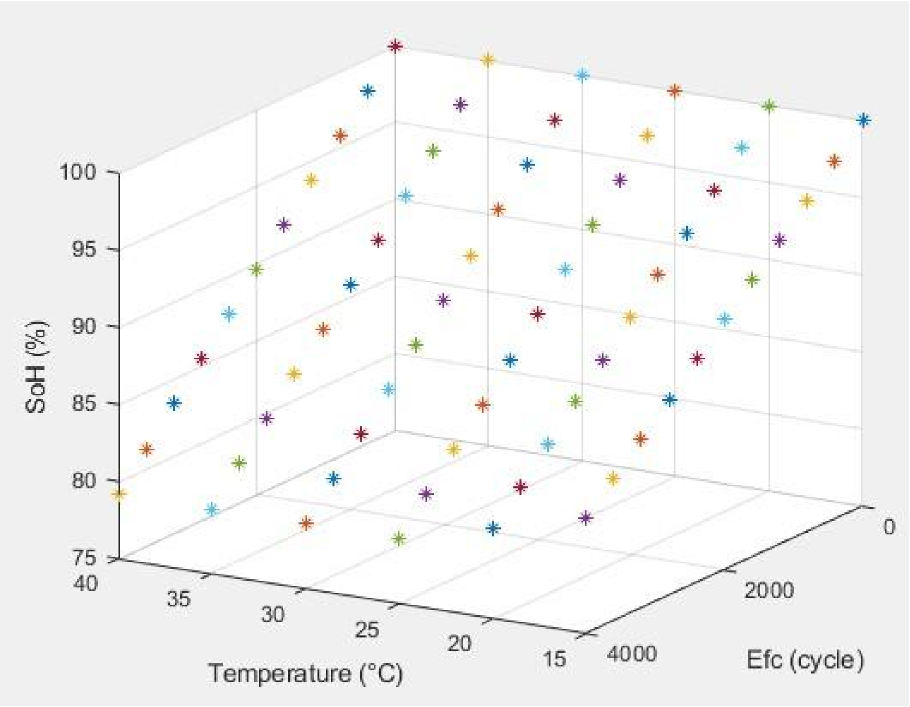


Figure 4.2.1.1: SoH reference points for each combination of temperature and Efc.

4.2.2 Coulomb efficiency computation

Given a fully charged/discharged battery which corresponding nominal capacity Qn, it needs Qn/C_rate hours to fully discharged/charged, however, it is not always true cause of the irreversibility of layered cathode materials during the first charge-discharge process [20]. The coulomb efficiency η has to be found in order to get the real time of fully charged/discharged the battery at a given temperature and aging condition, so at this point a 1C-rate constant discharged test is performed.
the formula for computing the Coulomb efficiency is derived accordingly to the performed choices during the tests (I(t) = Qn ∗ 1 C, t>=0 and SoC(0) = 100%, T = Ti , Ef c = Ef cj).

$$SoC(t) = SoC(0) - \frac{\eta(T,Efc)}{Q_n}\int_0^t Q_n * 1C d\tau$$

$$SoC(t) = SoC(0) - \eta(T,Efc)\int_0^t d\tau$$

$$SoC(t) = SoC(0) - \eta(T,Efc)t$$

$$SoC(t_{end}) = SoC(0) - \eta t_{end}$$

$$\eta(T,Efc) = \frac{SoC(0) - SoC(t_{end})}{t_{end}}$$

Knowing the relationship between Efc and SoH from the previous paragraph, the obtained values of η are fitted with a one-degree polynomial in the SoH variable for each temperature reference, by using Least Square. The obtained results are the following:

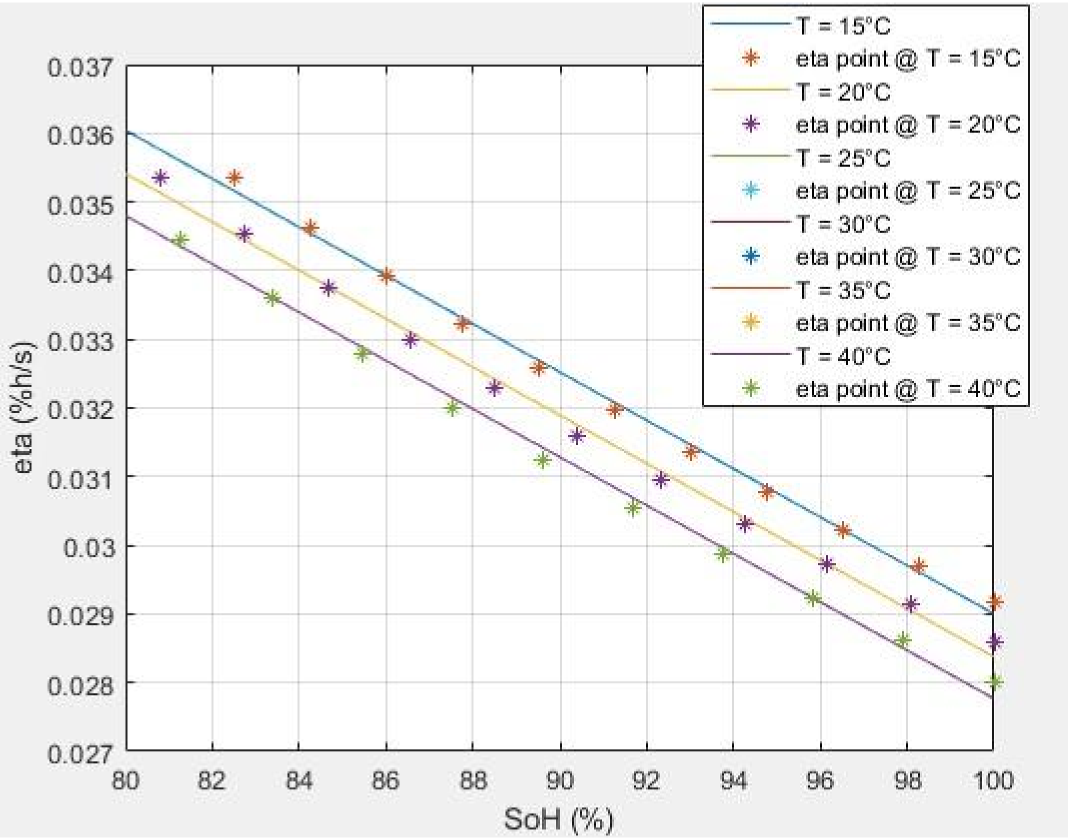


Figure 4.6: Coulomb efficiency points and fitted polynomials for different temperature

4.2.3 Open Circuit Voltage curve identification

The open circuit voltage of the battery is critical to accurate estimate the states of the battery, in this case, a pulse discharge test (PDT) is performed for each given Efc and temperature, the resolution SoC is chosen 5%, the goal is to find open circuit voltage for each SoC

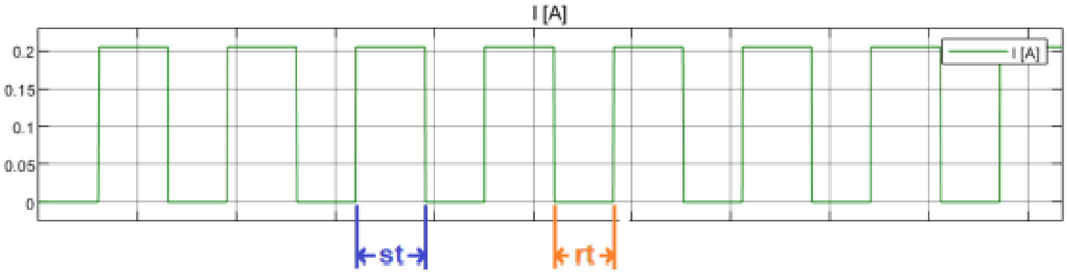


Figure 4.7: Example of a Pulse discharge current profiles. st: stimuli time, rt: rest time.

Note that rest time should be held long enough to allow the chemical reaction inside the battery to equilibrate, in our case is 30 minutes

As an example, The following images refers to one of the performed tests, which is relative to 25 ◦C ambient temperature, Efc equal to 1600 (corresponding to 91.68% of SoH):

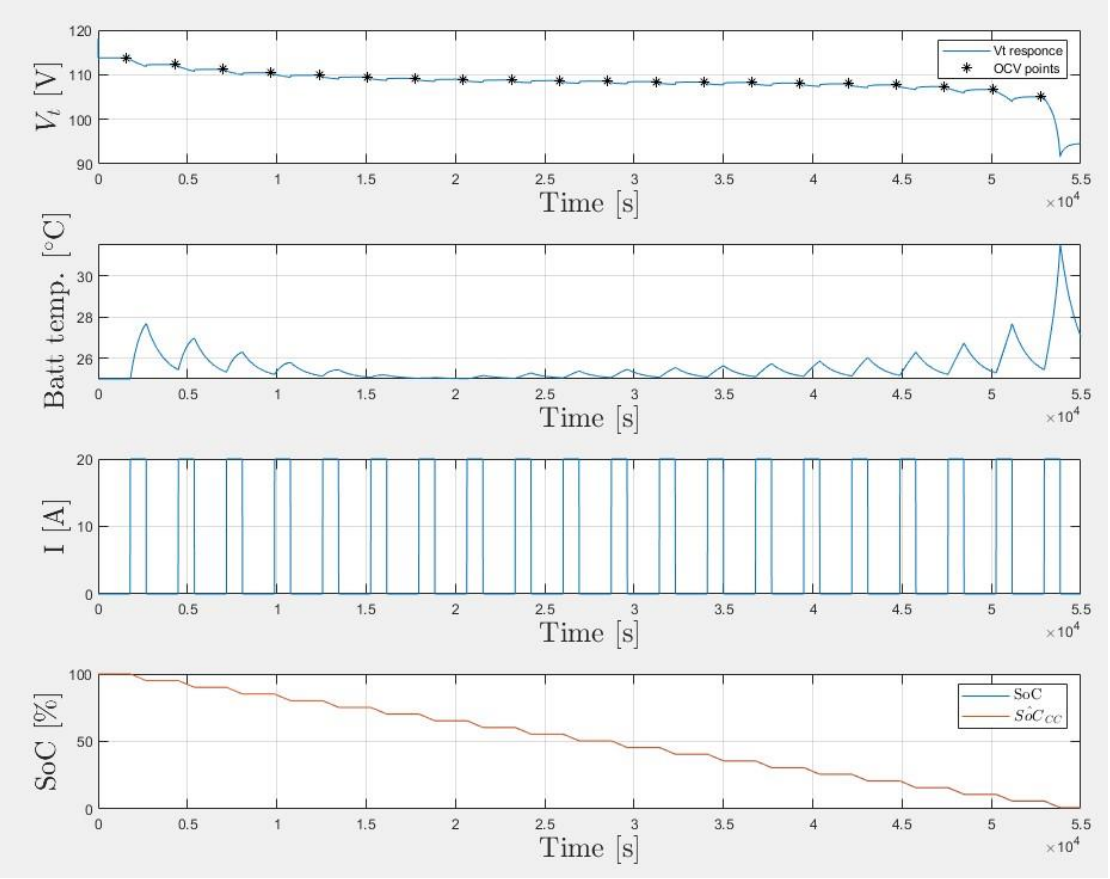


Figure 4.8: The acquired data during PDT.

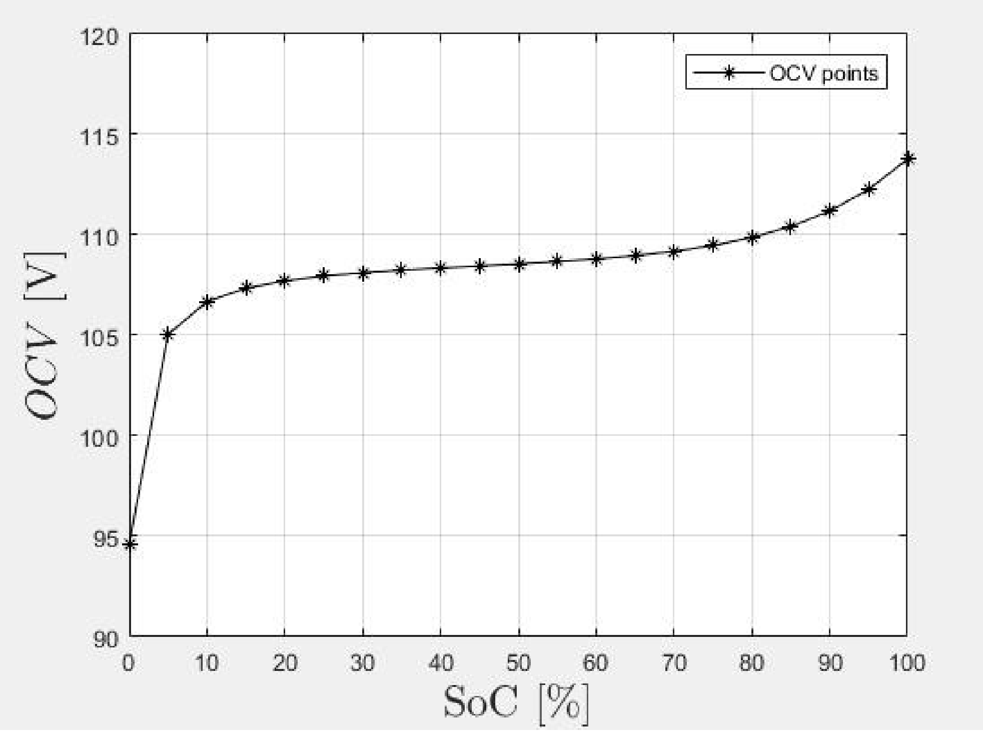


Figure 4.9: Linearly interpolated OCV data points vs SoC percentage.

The obtained OCV data points at different temperature, Efc, and SoC are used to find a two variable polynomial surface for different temperature. The chosen degree is 11 and it has been found using "polyfitn" function

[21]

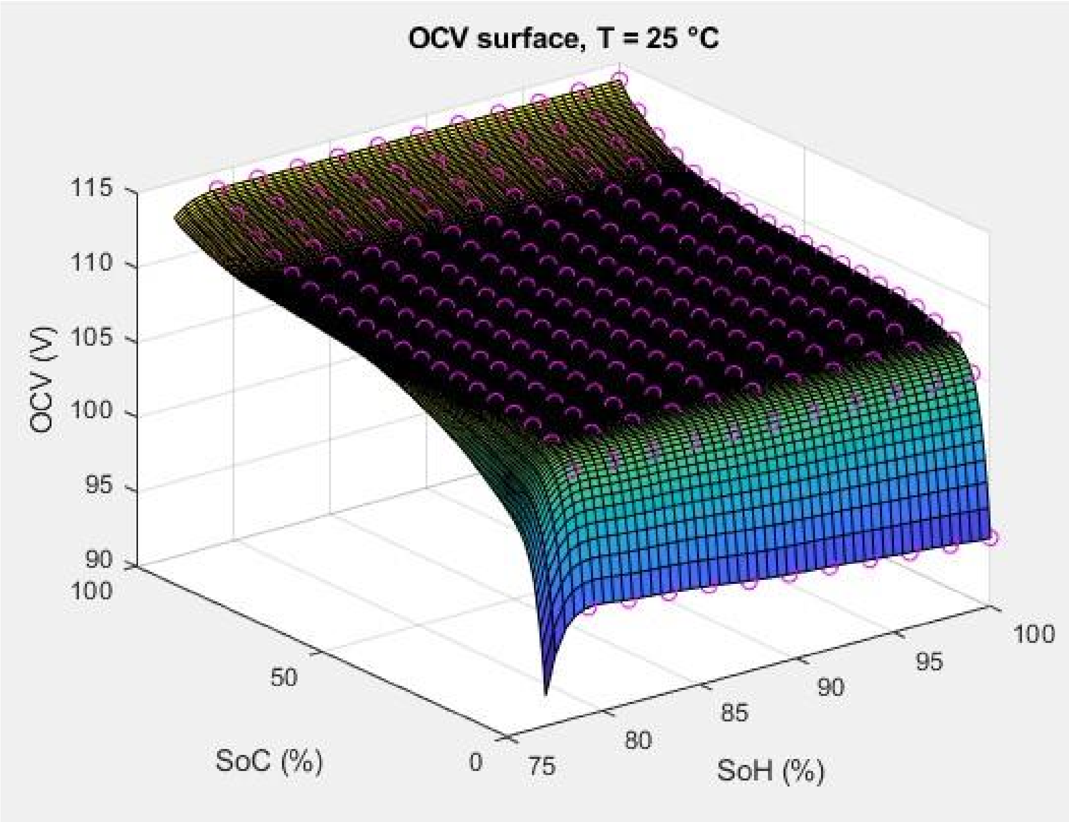


Figure 4.10: The OCV points and the relative surfaces are assumed to be constant in the temperature range [Tamb, Tamb + 5] 25◦C

The obtained OCV points are compared with the ones provided by the cell manufacturer:

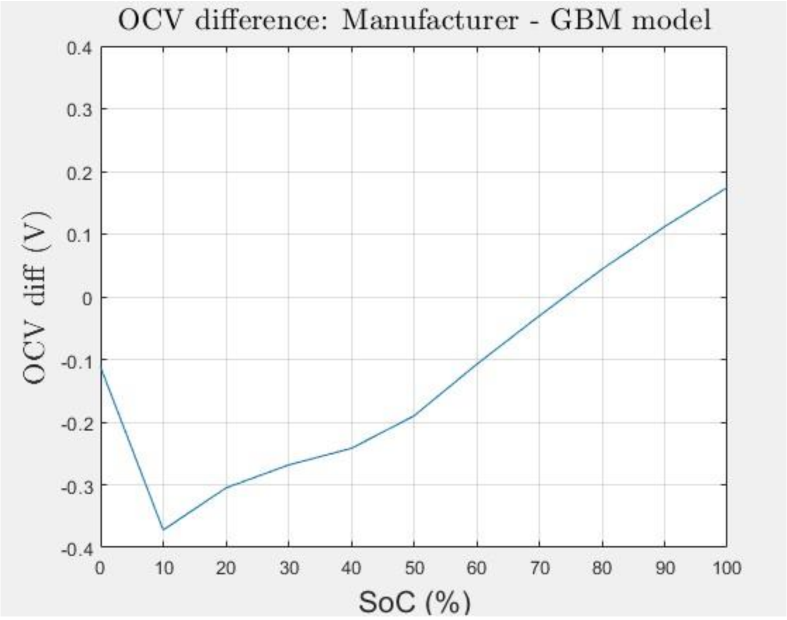


Figure 4.11: Difference between identified OCV data points and the ones provided by the cell manufacturer at BOL condition, temperature of 25°C and 10 % SoC resolution.

4.2.4 Equivalent Circuit Model identification

As mentioned in paragraph 4.2, the First Order Thévenin Model is the chosen ECM. The system of equations that describe its behavior

$$FOTM: \begin{cases} C\dfrac{dV_1}{dt} + \dfrac{V_1}{R_1} = I \\ V_T = OCV(SoC) - V_1 - R_{int}I \end{cases}$$

By applying Laplace transform, the first equation of FOTM becomes:

$$V_1 = \frac{I}{C(s + \frac{1}{R_1 C})}$$

Substituting in the second equation of FOTM:

$$V_D = V_T - OCV(SoC) = -\frac{I}{C(s + \frac{1}{R_1 C})} - R_{int}I$$

Then, the transfer function is:

$$\frac{V_D}{I} = -\frac{R_{int}s + \frac{R_{int}+R_1}{CR_1}}{s + \frac{1}{R_1 C}}$$

By applying the forward rule and passing to z domain (s = (z-1)/Δt) the above equation can be rewritten as:

$$\frac{z-1}{\Delta t}V_D + \frac{V_D}{R_1C} = -\frac{R_{int}}{\Delta t}(z-1)I - \frac{R_{int}+R_1}{R_1C}I$$

Where k is a discrete time instant, Δt sampling interval and τ = R1C. By changing variable and parametrizing :

$$V_D(k) + a_1V_D(k-1) = b_1I(k) + b_2I(K-1)$$

where:

$$\begin{cases} a_1 = \dfrac{\Delta t}{\tau} - 1 \\ b_1 = -R_{int} \\ b_2 = R_{int} - \dfrac{\Delta t}{\tau}(R_{int}+R_1) \end{cases}$$

Equation can be rearranged as following:

$$\begin{aligned} V_D(k) &= -a_1V_D(k-1) = b_1I(k) + b_2I(K-1) \\ y(k) &= [-V_D(k-1)I(k-1)I(k)][a_1b_1b_2]^T \\ y &= A * \theta \end{aligned}$$

The function above can be used along with Least square algorithm in order to identify parameters a1, b1, b2 and consequently Rint, R1, C, by using N sampled data.

$$\theta = (A^T A)^{-1} A^T y$$

$$\begin{cases} \tau = \dfrac{\Delta t}{a_1 + 1} \\ R_{int} = -b_1 \\ R_1 = (R_{int} - b_2)\dfrac{\tau}{\Delta t} - R_{int} \\ C = \dfrac{\tau}{R_1} \end{cases}$$

The ECM parameters RINT ,R1, C and τ are depending on battery average temperature, SoH and SoC. In order to find their values for each combination of temperature, SoH and SoC, 1C-rate Pulse Discharge Test is applied. Similarly, to the previous tests, the temperature and Efc are chosen according to 4.1, while SoC goes from 100 % to 0 % with a resolution of 5 %. Before applying the test, the battery is in thermal equilibrium with a certain ambient temperature, start from its full-charged state and at a specific aged status (Efc).

The necessary data consist in current, terminal voltage samples (acquired at 1Hz) and OCV which is find at the previous paragraph. Each SoC percentage value is associated with a batch of the acquired data and the relative OCV. Batches of data are chosen in the following way: The first batch at 100 % SoC corresponds to the discharge part of the voltage response, while all the other batches at different SoC level, correspond to the raising part of the voltage response. The first batch starts rt/9 seconds before the current raising part and ends when the current pulse ends. The other batches start on the falling part of the current pulse and ends after 8/9 rt seconds.

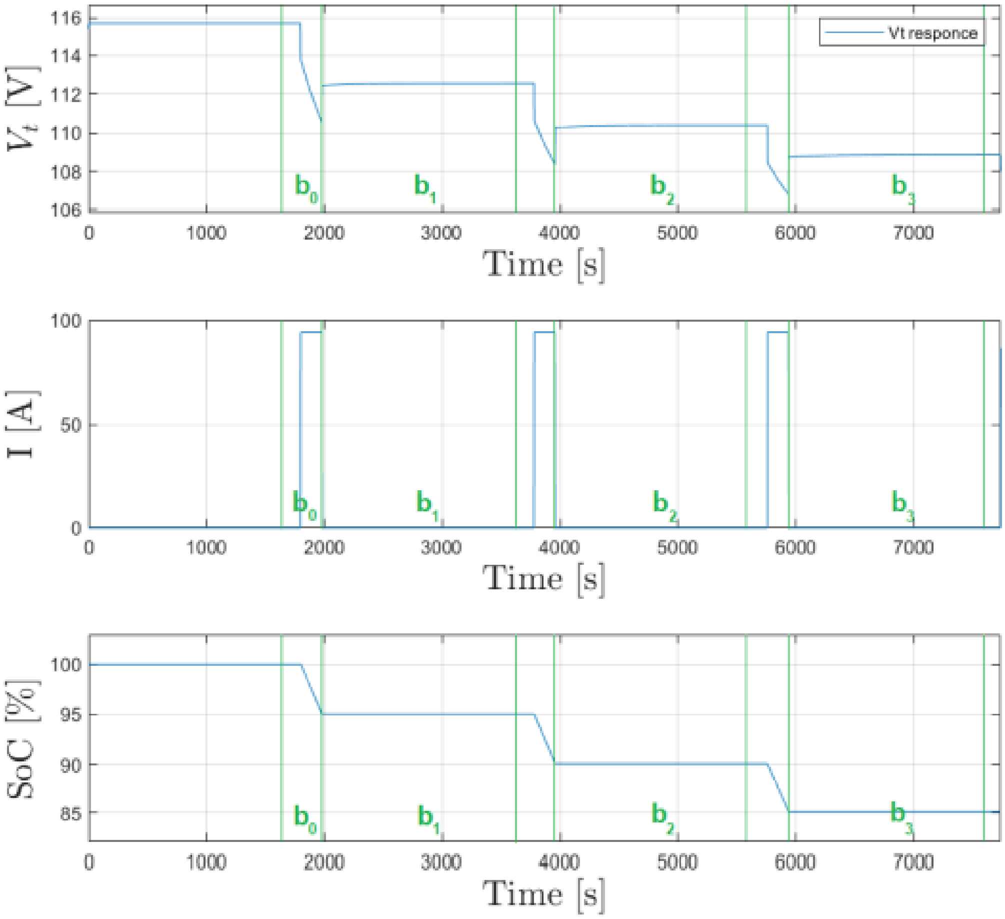


Figure 4.12: Example of batches partitioning for SoC interval [85, 100] %. The signals refers to a 1C PDT at 25◦C and 100% SOH.

Inside each batch, the voltage response, current stimuli and the computed OCV are used along with equations 4.14, 4.15 and 4.16 to find the ECM parameters in the relative context.

The obtained parameters at different temperature, Efc, and SoC are used to find a two variable polynomial surface for different temperature by using "polyfitn" function [21]. The chosen polynomial degrees are: 8 for RINT , 11 for R2, 2 for C and 8 for τ .

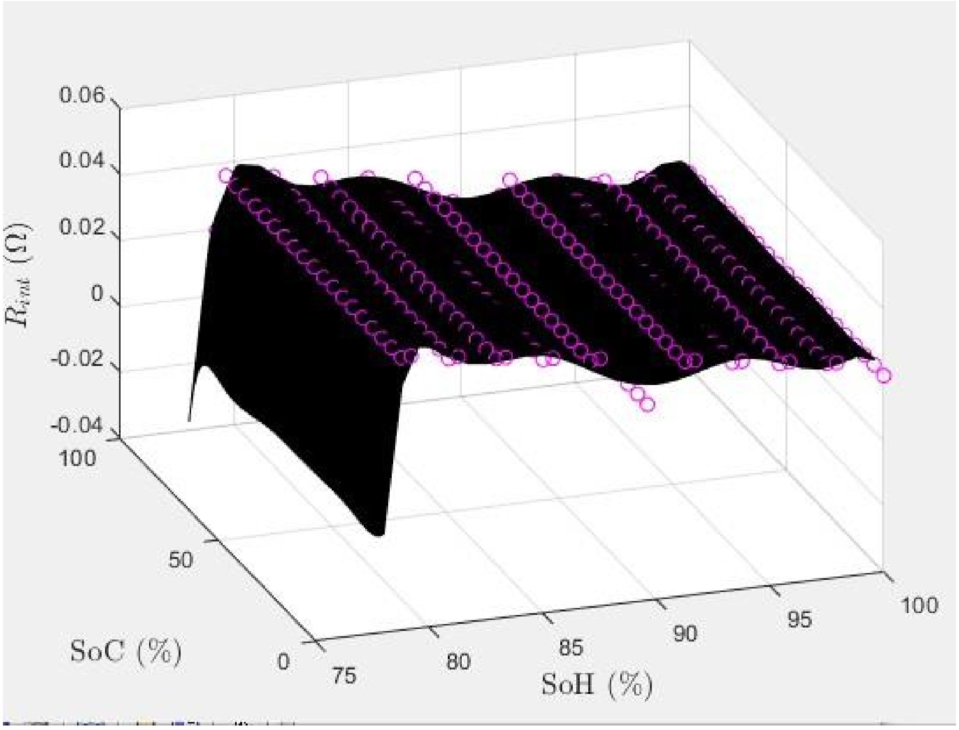


Figure 4.13: The RINT points and the relative surface at 25◦C.

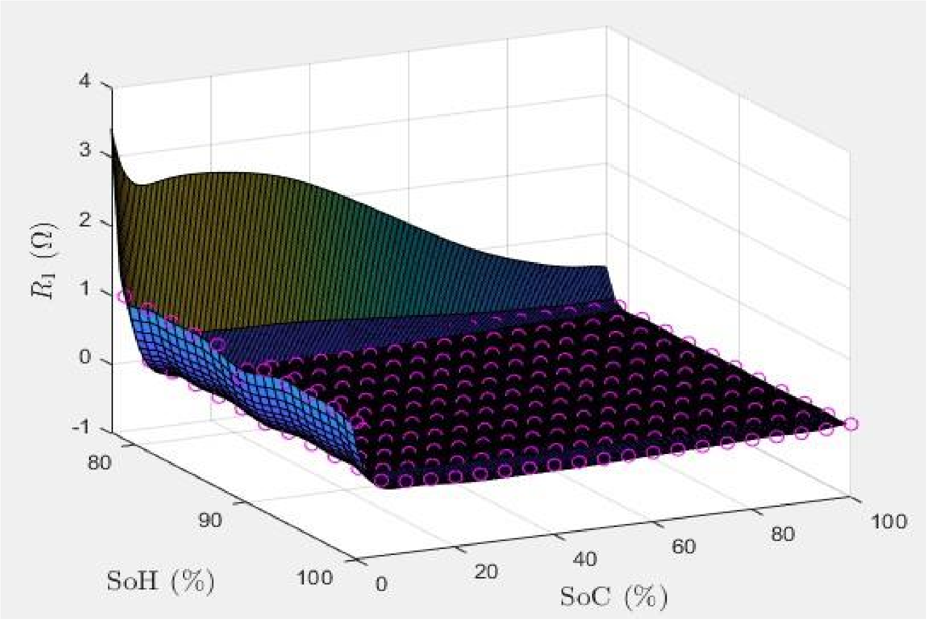


Figure 4.14: The R1 points and the relative surface at 25◦C.

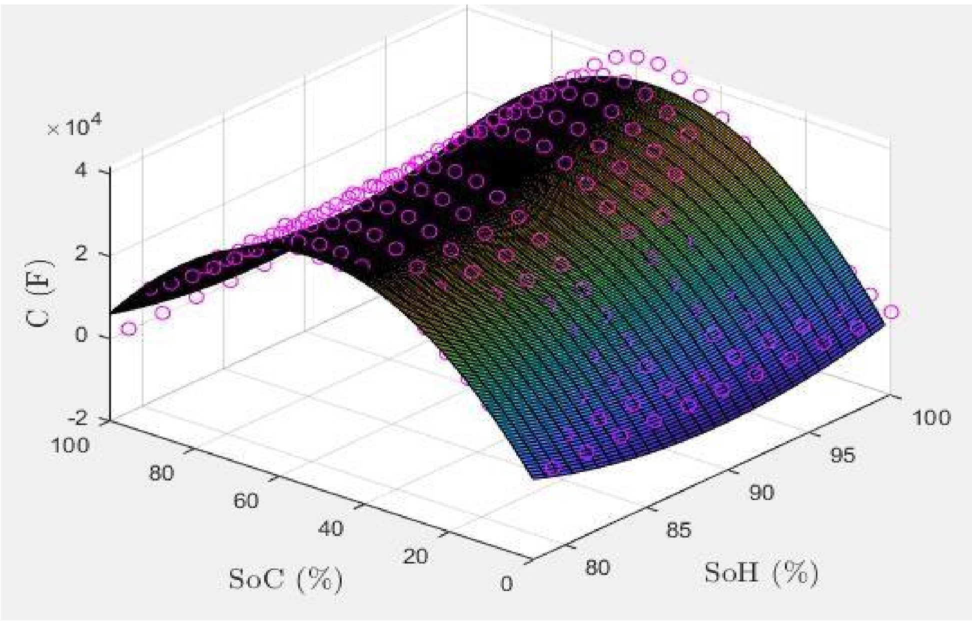


Figure 4.15: The C points and the relative surface at 25◦C.

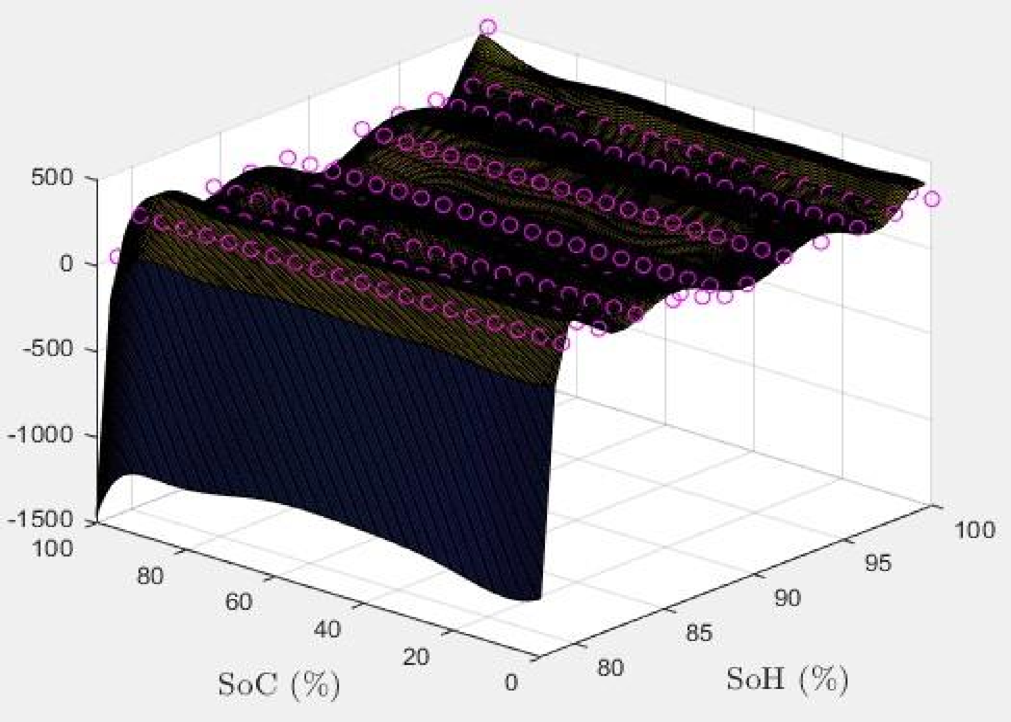


Figure 4.16: The τ points and the relative surface at 25◦C.

The parameters data are assumed to be constant in the temperature range

[Tamb, Tamb + 5].

4.2.5 Equivalent Circuit Model comparison

At this point a comparison between GBM and FOTM model is made in order to evaluate the goodness of the GBM model, the method is to implement 1C-rate CDT, and take the corresponding battery terminal voltage Vt as the reference.

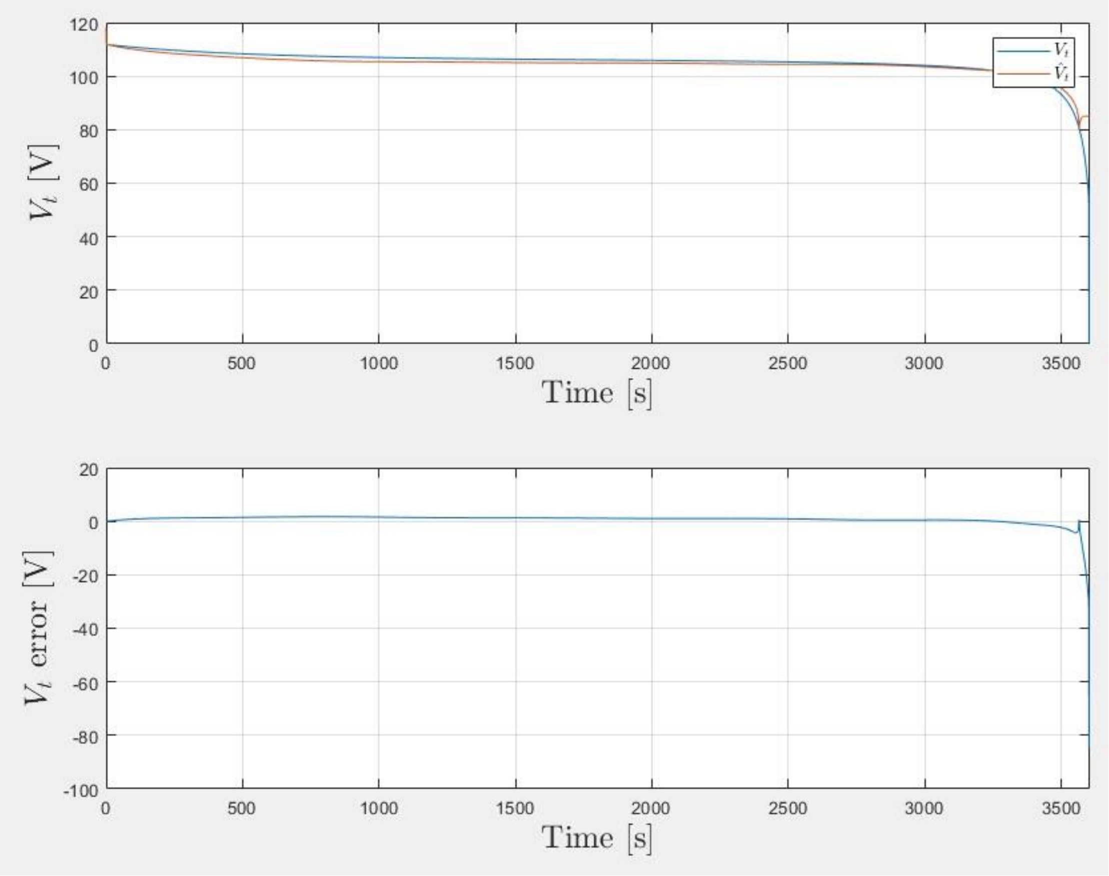


Figure 4.17: Comparison between FOTM and GBM under 1C CDT @ 25 ◦C and 100% SoH.

The voltage response of the ECM is pretty like the one produced by the GBM but becoming worse when SoC is near to 0. The average error is 0.6082 V, minimum error -84.6913 (SoC=0%) and the maximum error is 1.6993V

Chapter 5 Extended Kalman Filter

In paragraph 3.2.2 the first principles estimation approaches based on ECMs are addressed and the theory of the Extended Kalman Filter is explained. In this chapter the EKF is considered as classical approach and is enhanced with OCVmap technique to better estimate SoC and SoH.

An important aspect of choosing an EKF was the system model we chose. During its development, the first-order Thévenin (FOTM) ECM model was considered because it offers a good compromise between complexity and accuracy in representing the electrical dynamic behavior of a given battery.

Recall the basic equations of the extended Kalman filter that we introduced in the previous chapters:

$$\begin{aligned} x_k &= f\left(x_{k-1}, u_k, \omega_k\right) \\ z_k &= h\left(x_k, \nu_k\right) \end{aligned}$$

The considered FOTM is described by the following system of equations:

$$FOTM: \begin{cases} \dfrac{dV_1}{dt} + \dfrac{V_1}{\tau_{RC}} = \dfrac{I}{C} \\ V_T = OCV - V_1 - R_{int} I \\ \tau_{RC} = R_1 C \end{cases}$$

According to the scenario we want to apply, the chosen system state vector is:

$$x_k = \begin{bmatrix} SoH_k \\ SoC_k \\ V_{1_k} \end{bmatrix}$$

The expression for SoH is obtained by modelling the decreasing behavior of the SoH in time according to the calendar life provided by the cell manufacturer:

$$SoH(t) = -\frac{20}{78840000} t + 100$$

where T is considered as sampling time. The expression for SoC is derived from the Coulomb Counter expression:

$$SoC(t) = SoC(t_0) - \frac{\eta}{Q_n} \int_{t_0}^{t} I(\tau) d\tau$$

Passing to discrete time the following substitution are performed:

$$\begin{cases} t = (k' + 1)T \\ t_0 = k'T \end{cases}$$

Where k' is a generic discrete time instant and T is the sampling time. Then coulomb counter expression become:

$$SoC((k' + 1)T) = SoC(k'T) - \frac{\eta}{Q_n} \int_{k'T}^{(k'+1)T} I(\tau) d\tau$$

By changing variable $\tau = k\ 0T + \sigma$, $0 \leqslant \sigma \leqslant T$, the previous eq. is rewritten as:

$$SoC((k' + 1)T) = SoC(k'T) - \frac{\eta}{Q_n} \int_{0}^{T} I(k'T + \sigma) d\sigma$$

Assuming a zero-order holder $I(k'T + \sigma) = I(k'T)$ and then:

$$SoC((k' + 1)T) = SoC(k'T) - \frac{\eta}{Q_n} \int_{0}^{T} I(k'T) d\sigma$$

Now we define K=K'T, And make SoC and SoH directly related by adding the dependency between SoH and SoC to the coulomb counter expression, the previous expression becomes:

$$SoC(k + 1) = SoC(k) - \frac{100\eta T}{SoH(k) Q_{BOL}} I(k)$$

Now we turn to the last state parameter: terminal voltage V1. From the above system of FOTM equations, we obtain the expression for V1k by integrating V1over the time period [t₀,t]:

$$V_1(t) = V_1(t_0)e^{\frac{t_0-t}{\tau_{RC}}} + \frac{e^{\frac{-t}{\tau_{RC}}}}{C}\int_{t_0}^{t} e^{\frac{\tau}{\tau_{RC}}} I(\tau)d\tau$$

Discretize it in the same way as before we get:

$$\begin{cases} t = (k'+1)T \\ t_0 = k'T \end{cases}$$

$$\tau = k'T + \sigma,\, 0 \leq \sigma \leq T$$

$$\Downarrow$$

$$V_1((k'+1)T) = V_1(k'T)e^{\frac{-T}{\tau_{RC}}} + \frac{e^{\frac{-(k'+1)T}{\tau_{RC}}}}{C}\int_{0}^{T} e^{\frac{k'T+\sigma}{\tau_{RC}}} I(k'T+\sigma)d\sigma$$

By Applying zero-order holder ,k = k’T and by performing some simplification:

$$\mathrm{I}(\mathrm{k'T} + \sigma) = \mathrm{I}(\mathrm{k'T})$$

$$V_1(k+1) = V_1(k)e^{\frac{-T}{\tau_{RC}}} + R_1(1 - e^{\frac{-T}{\tau_{RC}}})I(k)$$

So passing to discrete time, equations of the FOTM model become:

$$FOTM : \begin{cases} V_{1_{k+1}} = V_{1_k}e^{\frac{-T}{\tau_{RC}}} + R_1(1 - e^{\frac{-T}{\tau_{RC}}})I_k \\ V_{T_k} = OCV(SoC_k) - V_{1_k} - R_{int}I_k \\ \tau_{RC} = R_1C \end{cases}$$

At this point by considering equations above the state space representation of the system is:

$$\begin{cases} \begin{bmatrix} SoH_{k+1} \\ SoC_{k+1} \\ V_{1_{k+1}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{\frac{-\Delta T}{\tau_{RC_k}}} \end{bmatrix} \begin{bmatrix} SoH_k \\ SoC_k \\ V_{1_k} \end{bmatrix} + \begin{bmatrix} -\frac{20\Delta T}{78840000 I_k} \\ -\frac{100\eta_k \Delta T}{SoH_k Q_{BOL}} \\ R_{1_k}(1 - e^{\frac{-\Delta T}{\tau_{RC_k}}}) \end{bmatrix} I_k + w_k \\ y_k = OCV_k - V_{1_k} - R_{int_k} I_k + v_k \end{cases}$$

The system is non-linear and Jacobian matrices Fk, Hk has to be computed:

$$\begin{cases} F_k = \frac{\partial f(x,u)}{\partial x}\Big|_{(x,u)=(\hat{x}_k,u_k)} \\ H_k = \frac{\partial h(x,u)}{\partial x}\Big|_{(x,u)=(x_k^p,u_k)} \end{cases}$$

According to state space representation of the system, the matrices Fk and Hk can be computed as:

$$F_k = \begin{bmatrix} 1 & 0 & 0 \\ f_{2,1} & 1 & 0 \\ f_{3,1} & f_{3,2} & e^{\frac{-\Delta T}{\tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)}} \end{bmatrix}$$

Where:

$$f_{2,1} = -\frac{100\Delta T I_k}{Q_{BOL}(T_{amb})} \frac{\frac{\partial \eta(T_k, S\hat{o}H_k)}{\partial SoH} S\hat{o}H_k - \eta(T_k, S\hat{o}H_k)}{S\hat{o}H_k^2}$$

$$\begin{aligned} f_{3,1} = {} & \Delta T V_{1_k} e^{-\frac{\Delta T}{\tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)}} \frac{\frac{\partial \tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)}{\partial SoH}}{\tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)^2} + \\ & + I_k \frac{\partial R_1(T_k, S\hat{o}H_k, S\hat{o}C_k)}{\partial SoH}(1 - e^{-\frac{\Delta T}{\tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)}}) + \\ & - \Delta T I_k R_1(T_k, S\hat{o}H_k, S\hat{o}C_k) \frac{\frac{\partial \tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)}{\partial SoH}}{\tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)^2} e^{-\frac{\Delta T}{\tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)}} \end{aligned}$$

$$\begin{aligned} f_{3,2} = {} & \Delta T V_{1_k} e^{-\frac{\Delta T}{\tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)}} \frac{\frac{\partial \tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)}{\partial SoC}}{\tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)^2} + \\ & + I_k \frac{\partial R_1(T_k, S\hat{o}H_k, S\hat{o}C_k)}{\partial SoC}(1 - e^{-\frac{\Delta T}{\tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)}}) + \\ & - \Delta T I_k R_1(T_k, S\hat{o}H_k, S\hat{o}C_k) \frac{\frac{\partial \tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)}{\partial SoC}}{\tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)^2} e^{-\frac{\Delta T}{\tau_{RC}(T_k, S\hat{o}H_k, S\hat{o}C_k)}} \end{aligned}$$

The matrix Hk is computed as:

$$H_k = \begin{bmatrix} h_1 & h_2 & -1 \end{bmatrix}$$

Where:

$$h_1 = \frac{\partial OCV(T_k, SoH_k^p, SoC_k^p)}{\partial SoH} - \frac{\partial V_1}{\partial SoH} - \frac{\partial R_{int}(T_k, SoH_k^p, SoC_k^p)}{\partial SoH} I_k$$

and

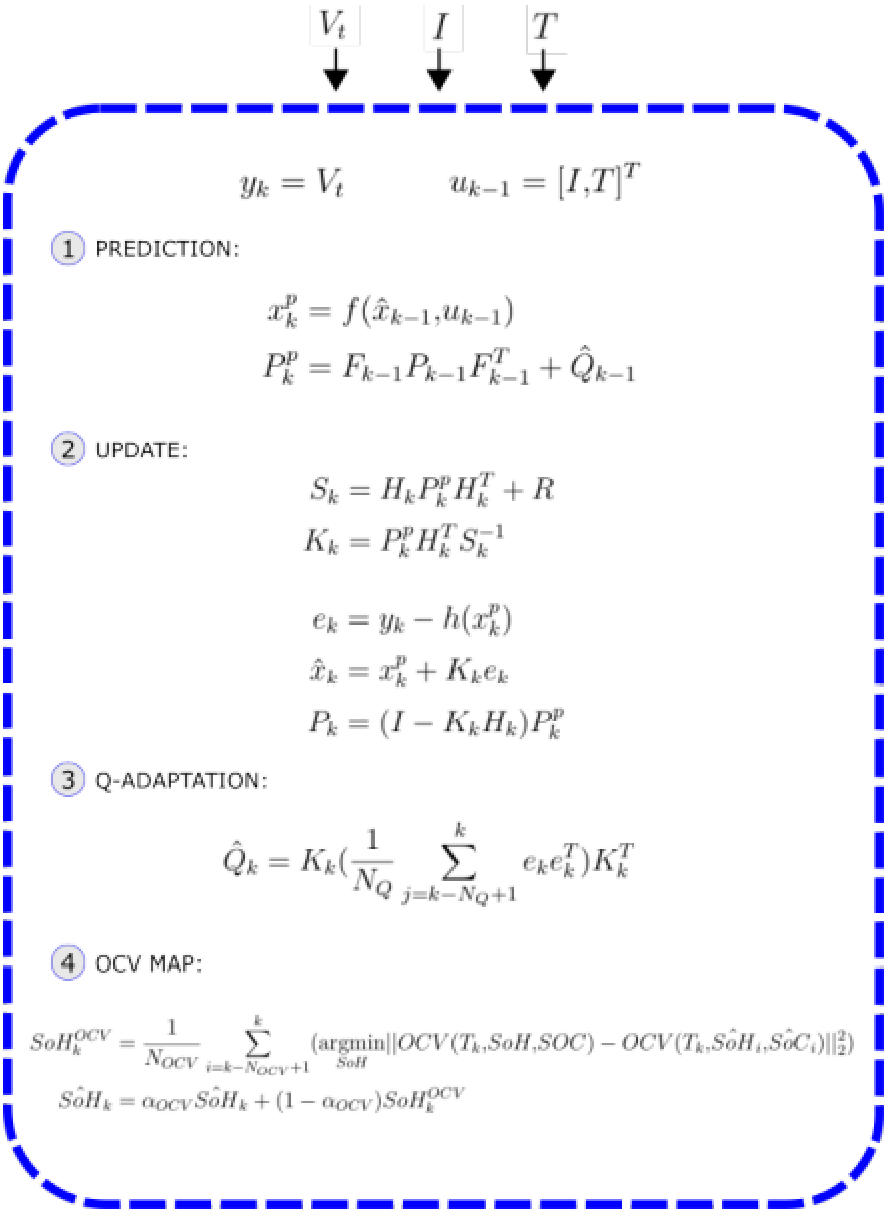
$$\begin{aligned}
\frac{\partial V_1}{\partial SoH} &= \frac{V_{1_{k+1}} - V_{1_k}}{SoH_{k+1}^p - SoH_k^p} = \\
&= \frac{V_1 e^{\frac{-T}{\tau_{RC}(T_k, SoH_k^p, SoH_k^p)}} + R_1(T_k, SoH_k^p, SoH_k^p)(1 - e^{\frac{-T}{\tau_{RC}(T_k, SoH_k^p, SoH_k^p)}})I_k - V_{1_k}}{(-\frac{-20\Delta T}{78840000} + SoH_k^p) - SoH_k^p} \\
&= -\frac{78840000}{20\Delta T}(1 - e^{\frac{-T}{\tau_{RC}(T_k, SoH_k^p, SoH_k^p)}})(R_1(T_k, SoH_k^p, SoH_k^p)I_k - V_{1_k})
\end{aligned}$$

$$h_2 = \frac{\partial OCV(T_k, SoH_k^p, SoC_k^p)}{\partial SoC} - \frac{\partial V_1}{\partial SoC} - \frac{\partial R_{int}(T_k, SoH_k^p, SoC_k^p)}{\partial SoC} I_k$$

and

$$\begin{aligned}
\frac{\partial V_1}{\partial SoC} &= \frac{V_{1_{k+1}} - V_{1_k}}{SoC_{k+1}^p - SoC_k^p} = \\
&= \frac{V_1 e^{\frac{-T}{\tau_{RC}(T_k, SoH_k^p, SoH_k^p)}} + R_1(1 - e^{\frac{-T}{\tau_{RC}(T_k, SoH_k^p, SoH_k^p)}})I_k - V_{1_k}}{(SoC_k^p - \frac{\eta(T_k, SoH_k^p)\Delta T}{Q_n} I_k) - SoC_k^p} \\
&= -\frac{Q_n}{\eta(T_k, SoH_k^p)\Delta T I_k}(1 - e^{\frac{-T}{\tau_{RC}(T_k, SoH_k^p, SoH_k^p)}})(R_1(T_k, SoH_k^p, SoH_k^p)I_k - V_{1_k})
\end{aligned}$$

To make the estimation of the predictive model more robust, we choose to adapt the EKF by iteratively updating the process noise covariance matrix (Q) based on the innovation term, which is also called AEKF. Furthermore, to enhance SoH and SoC estimation, on top of filtered state estimation, OCVmap technique is applied together with EKF. The figure below shows the overall algorithm:



V_t I T

$$y_k = V_t \qquad u_{k-1} = [I,T]^T$$

1 PREDICTION:

$$x_k^p = f(\hat{x}_{k-1},u_{k-1})$$
$$P_k^p = F_{k-1}P_{k-1}F_{k-1}^T + \hat{Q}_{k-1}$$

2 UPDATE:

$$S_k = H_k P_k^p H_k^T + R$$
$$K_k = P_k^p H_k^T S_k^{-1}$$

$$e_k = y_k - h(x_k^p)$$
$$\hat{x}_k = x_k^p + K_k e_k$$
$$P_k = (I - K_k H_k)P_k^p$$

3 Q-ADAPTATION:

$$\hat{Q}_k = K_k(\frac{1}{N_Q}\sum_{j=k-N_Q+1}^{k} e_k e_k^T)K_k^T$$

4 OCV MAP:

$$SoH_k^{OCV} = \frac{1}{N_{OCV}}\sum_{i=k-N_{OCV}+1}^{k}(\underset{SoH}{\text{argmin}}||OCV(T_k,SoH,SOC) - OCV(T_k,S\hat{o}H_i,S\hat{o}C_i)||_2^2)$$
$$S\hat{o}H_k = \alpha_{OCV}S\hat{o}H_k + (1-\alpha_{OCV})SoH_k^{OCV}$$

Figure 5.1: Overview of the EKF estimation algorithm for a k-esime iteration.

The first two set of equations illustrated in figure above represent the usual prediction and update step of the EKF, while the equation in step 3 implements a Q-adaptation based on a sliding window of size NQ. Step 4 implements the OCV map strategy to update the filtered version of the SoH: using a sliding window of size NOCV , the SoH obtained in correspondence of the minimum OCV error (computed by comparing the a-priori OCV data points and the computed OCV using filtered state) are collected; so at a given time the filtered value of the SoH is substituted with the exponential mean of the current filtered SoH and the mean value obtained through OCV map.

Initial values of the windows for both step 3 and 4 are the zero vector and their sizes (NQ, NOCV) along with αOCV become parameters to be tuned. Step 3 and step 4 are introduced to make more robust the estimation for both SoH and SoC [1].

5.1 Initial conditions and EKF tuning

The choice of initial state data is particularly important for the accuracy of predictions. So we assume system is supposed to start from a full charged state and after a long time is passed from its last usage, while its SoH is supposed to be unknown. The estimate of the initial condition is chosen randomly:

$$\hat{x}_0 = \hat{x}(k=0) = \begin{bmatrix} S\hat{o}H_0 \\ S\hat{o}C_0 \\ \hat{V}_{1_0} \end{bmatrix} = \begin{bmatrix} 91.9459 \\ 95.6547 \\ 0 \end{bmatrix} \begin{matrix} \% \\ \% \\ V \end{matrix}$$

Based on the premise of using Kalman filter, the true system initial condition is unknown and is modelled as a multivariate gaussian distributed random vector:

$$x_0 = x(k=0) \sim \mathcal{N}(\mu_{x0}, \Sigma_{x_0})$$

Although the initial condition is unknown, some assumption on its statistical parameters can be done, that is, set a reasonable falling range for random data. The SoH is supposed to belong to the interval [80, 100] % with 99.73% of probability. This can be traduced in the following interval [μSoH0 − 3σ SoH0 ,μSoH0 + 3σSoH0], where μSoH0 is 90 % and σSoH0 = 10/3%. SoC is supposed to belong to an interval equal to [μSoC0 − 3σSoC0 ,μSoC0 + 3σSoC0], where μSoC0 is 95 % and σSoC0 = 5/3%. In this way, with 99.73% of probability, the SoC initial condition of the system belong to the interval [90, 100] %. Similarly to the previous cases, V10 is supposed to be near to 0. Its variation belong to the interval [μV10 − 3σV10 ,μV10 + 3σV10] , where μV10 is 0.1 V and σV10 = 0.1/3 V.

So the statistical parameters attributed to the state initial condition are known by assumption:

$$\mu_0 = \begin{bmatrix} 90 \\ 95 \\ 0.1 \end{bmatrix} \begin{matrix} \% \\ \% \\ V \end{matrix}$$

$$\begin{aligned}
\Sigma_{x0} &= E\left[(x_0 - E[x_0])(x_0 - E[x_0])^T\right] = \\
&= E\left[(x_0 - \mu_{x0})(x_0 - \mu_{x0})^T\right] = \\
&= \begin{bmatrix} \sigma^2_{SoH_0} & \sigma_{SoH_0SoC_0} & \sigma_{SoH_0V_{1_0}} \\ \sigma_{SoC_0SoH_0} & \sigma^2_{SoC_0} & \sigma_{SoC_0V_{1_0}} \\ \sigma_{V_{1_0}SoH_0} & \sigma_{V_{1_0}SoC_0} & \sigma^2_{V_{1_0}} \end{bmatrix} = \\
&= \begin{bmatrix} \sigma^2_{SoC_0} & \rho_{HC}\sigma_{SoH_0}\sigma_{SoC_0} & \rho_{HV}\sigma_{SoH_0}\sigma_{V_{1_0}} \\ \rho_{HC}\sigma_{SoC_0}\sigma_{SoH_0} & \sigma^2_{SoC_0} & \rho_{CV}\sigma_{SoC_0}\sigma_{V_{1_0}} \\ \rho_{HV}\sigma_{V_{1_0}}\sigma_{SoH_0} & \rho_{CV}\sigma_{V_{1_0}}\sigma_{SoC_0} & \sigma^2_{V_{1_0}} \end{bmatrix}
\end{aligned}$$

Where the parameters ρHC,ρHV and ρCV are the correlation coefficients and are found by trial and error approach. The EKF matrices R, Q0 and P0 are chosen in the following way:

$$R = E[(\tilde{y}_k - E[\tilde{y}_k])(\tilde{y}_k - E[\tilde{y}_k])T],\ \tilde{y} = y_k - \hat{y}_k$$

Which is chosen according to the statistic properties of the measurement error which is described next.

$$P_0 = E[(\tilde{x}_0 - E[\tilde{x}_0])(x_0 - E[x_0])^T], \tilde{x}_0 = x_0 - \hat{x}_0$$

According to the Simulink schema, the WLTP cycle test is performed at 25 ◦C and for all the SoH reference conditions {100, 97.92, 95.84, 93.76, 91.68, 89.6, 87.52, 85.44, 83.36, 81.28, 79.20} ％ on the dynamic simulator. The sampled (1Hz) acquired data in terms of battery voltage, current, and temperature are then corrupted by measurement random noise uniformly distributed, in order to introduce disturbances by emulating measurement sensors. The characteristics of the noises for each measurement are the following:

- max(wV) = 0.3 V, min(wV) = -0.3 V, µV = 0

- max(wI) = 0.05 A, min(wI) = -0.05 A, µI = 0

- max(wT) = 1 ◦C, min(wT) = ◦C, µT = 0

Referring to the scenario with 100% SoH, the picture attached here is used to find the measurement error of the FOTM model and then a suitable value for the R matrix of the EKF.

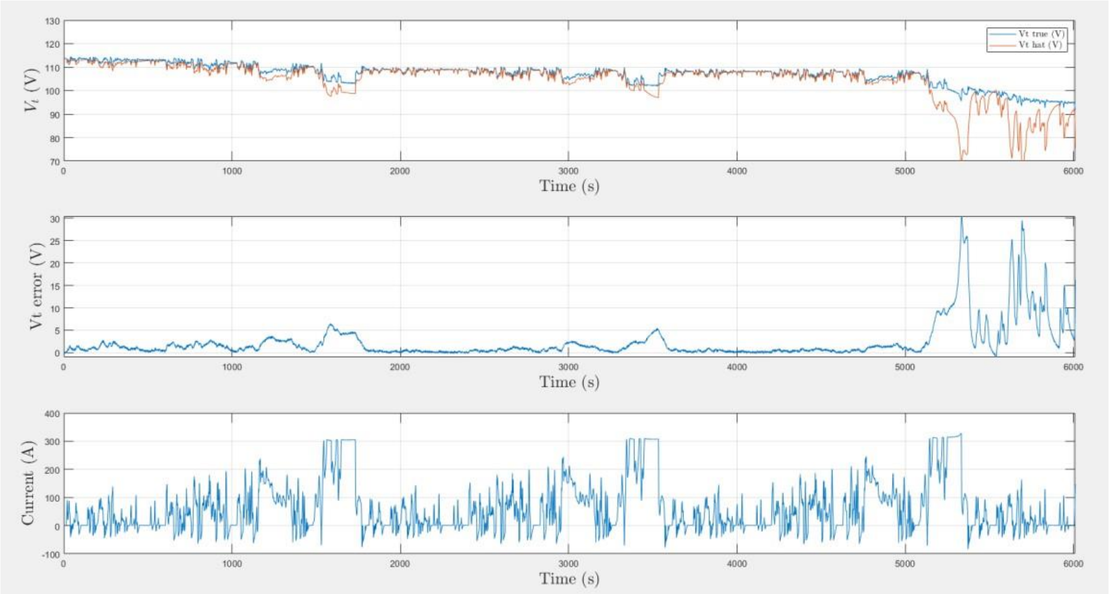


Figure 5.1.1: Signals obtained by comparing data produced by the dynamic simulator and the FOTM model by performing wltp at 25 ◦C ambient temperature and 100% SoH.

Referring to the previous figure, the standard deviation of the measurement error is $\sigma_{\tilde{y}}$ = 4.13269 (V), then the matrix R is equal to $\sigma_{\tilde{y}}^{2}$, In the next chapter we will detail the results of the test.

Chapter 6 test results

The previous chapters have introduced various methods for estimating SoC, SoH of the battery and the settings of Kalman filter. This chapter will focus on testing the performance of Kalman filter algorithm and the comparison with neural network methods.

6.1 tests in virtual test bench

Regarding the introduction of virtual test bench and its completion in chapter 4.1, this paragraph will focus on using this virtual test bench to simulate the driving situation of an electric vehicle to complete the tests of our extended Kalman filter algorithm.

The following tests are done in the WLTP driving situation, the ambient temperature is 25∘C, the equivalent full cycles (Efc) of the battery is [0,800,2000;4000], which the corresponding SoH is [100%,95.8%,89.6%,79.2%]:

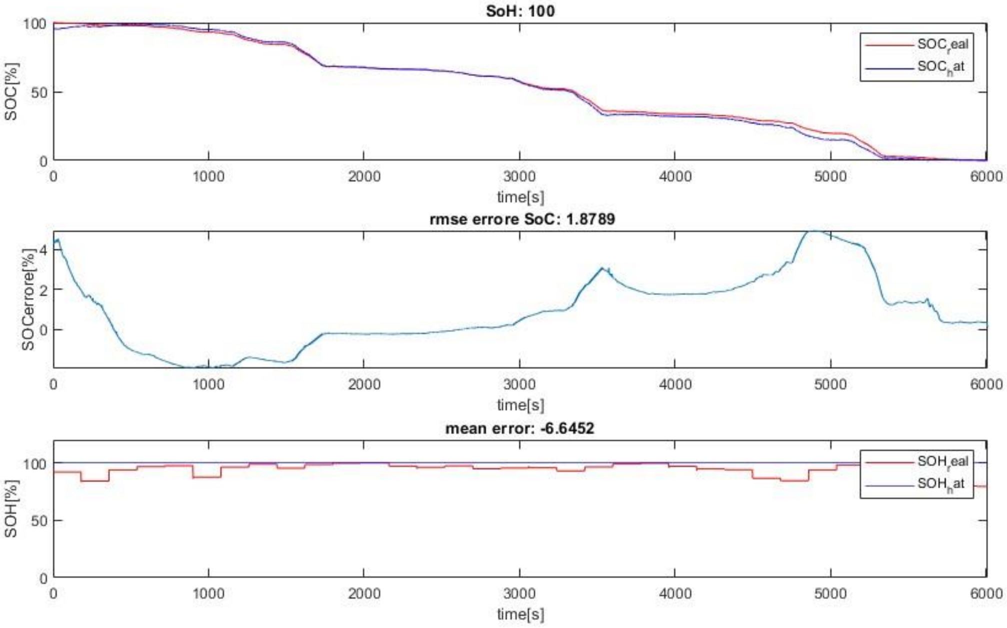


Figure 6.1.1 Example of test results: Efc=0, SoH=100%

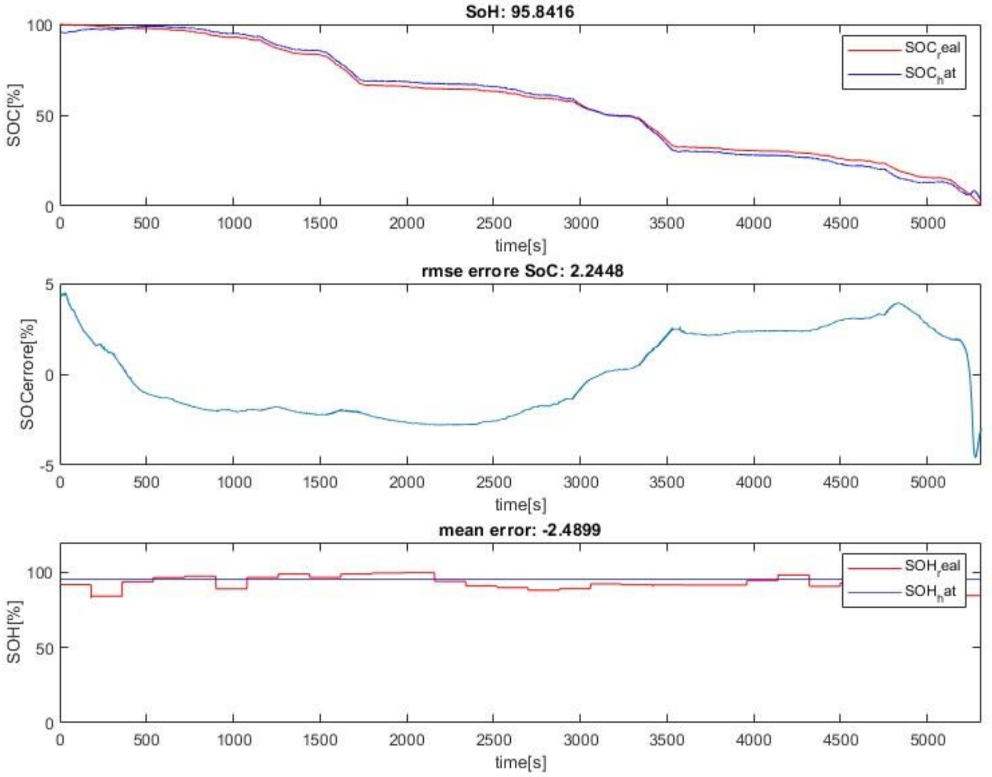


Figure 6.1.2 Example of test results: Efc=800 SoH=95.84%

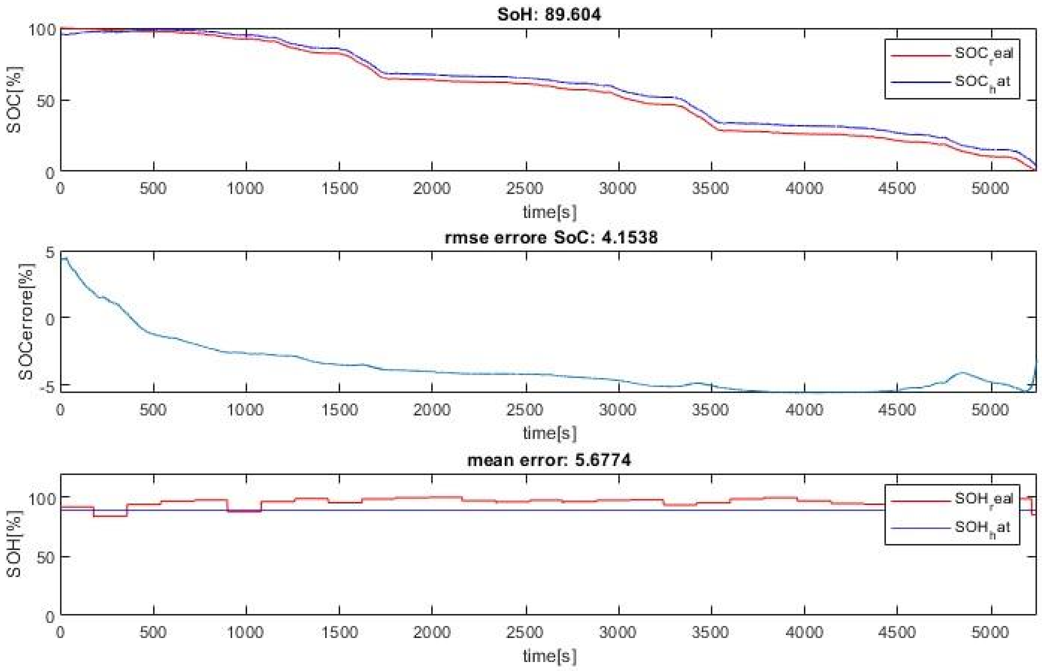


Figure 6.1.3 Example of test results: Efc=2000, SoH=89.6

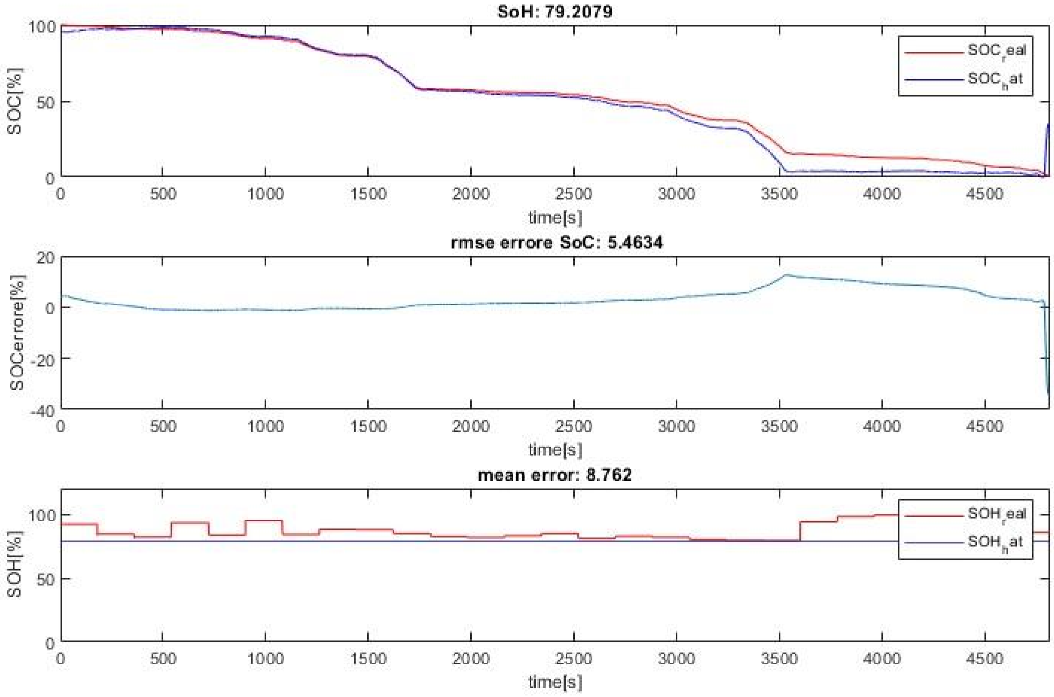


Figure 6.1.4 Example of test results: Efc=4000, SoH=79.2

As can be observed by looking at the results, when the estimate of the SoH approaches the truth one the error in the SoC estimation reduces thanks to the adaptivity of the algorithm. Even if the starting initial condition of the state estimate is different from the truth one, the SoC estimate tends to the reference in a finite time, which is a very important characteristic. The average mean error and root mean square error for SoC estimate over all the tests is under 3 %. It is observed that the SoH estimate fluctuates near the reference but is not always accurate. This can be due to the presence of the noise in the measurements and the precision of the OCV curve that can penalize the SoH update in time. The global average error of SOH can also be controlled within 5%, but its fluctuation range is relatively severe, and the prediction stability is not as good as that of SOC.

The effectiveness of the algorithm in the virtual environment test bench has been verified in the previous chapter, in this photograph there will be more tests to verify the robustness of the algorithm, which means that the algorithm can maintain good performance under any interference and different road conditions.

1. Worldwide harmonized Light vehicles Test Procedure (WLTP): a chassis dynamometer test cycle for the determination of emissions and fuel consumption from light-duty vehicles

Results by using WLTP class 3 driving cycle, Tambient=25◦C:

Efc	SOC mean error %	SOH mean error %
0	0.72	5.7
400*	1.63	6.9
800	0.05	2.38
1200	-1.5	1.2
1600	3.09	2.95
2000	3.5	5.6
2400	-6.5	8.26
2800	4.82	7.7
3200	4.81	2.93
3600	4.1	5.1
4000	3.15	8.52

Average SOC error： 3.08%, average SOH error： 5.2%

2. NEDC: used as reference cycle for homologating vehicles until Euro6 norm in Europe and some other countries.

Results by using NEDC class 3 driving cycle, Tambient=25◦C:

efc	SOC mean error %	SOH mean error %
0	6.1	9.3
400*	5.7	8.8
800	4.8	7.1
1200	2	2.9
1600	1.5	0.96
2000	1.58	0.57
2400	0.27	3.8
2800	1	6.2
3200	1.97	8.1
3600	2.8	10.5
4000	5	12

Average SOC error：2.97 %, average SOH error：5.38%

3. FTP-75: The FTP-75 (Federal Test Procedure) has been used for emission certification and fuel economy testing of light-duty vehicles in the United States

Results by using FTP-75 class 3 driving cycle, Tambient=25◦C

efc	SOC mean error %	SOH mean error %
0*	7.64	-10
400	6.2	5.2
800	2.8	5.6
1200	5.3	2.8
1600	5.94	0.06
2000	1.56	3.35
2400	2.42	3.36
2800	1.4	5.25
3200	0.14	6.3
3600	1.63	8.8
4000	1.06	9.57

Average SOC error：3.28 %, average SOH error：5.48%

Since the vehicle may be driven in any season, it is equally important to test for different ambient temperatures:

Results by using WLTP 3 driving cycle, Equivalent Full Cycles (EFC) equal to 800, Tambient=[15 20 25 30 35 40]◦C

Average SOC error：0.706 %, average SOH error：2.17%

T◦C	15	20	25	30	35	40
SOC mean error %	1.25	0.75	0.05	0.86	0.08	1.25
SOH mean error %	2.83	1.16	2.38	4.4	1.51	0.79

It can observed that in each driving condition (WLTP NEDC FTP-75) our EKF algorithm maintains high accuracy in estimating SoC, which corresponding average error is : 3.08%, 2.97%, 3.28%, as well as in different ambient temperature cases, meanwhile higher estimation error for SoH, this is mainly cause of the precision of the OCV curve, but the most important conclusion is that our algorithm has a very limited difference in the results under different driving condition and temperatures, which means it has high robustness, at least in the virtual test bench.

6.2 Comparison between Extended Kalman filter and Neutral network

The effectiveness and robustness of the extended Kalman filter algorithm have been demonstrated in the virtual environment in the last chapter. This chapter will focus on comparing the performance of Extended Kalman Filter and Neutral Network method. As described in chapter 3.3.2, Artificial intelligence is also an emerging and promising method to measure the states of the battery, so it will be interesting to compare the performance of these two techniques.

This part of the work was done in collaboration with another colleague MinNing, who was focused on the design and training of neural networks and provided me with the algorithm in black box mode, and I was responsible for providing training data and do the further tests.

The data used to train the neural network is performed in virtual test bench, the driving condition is WLTP, and the corresponding SoH of the battery is 100%, 89.6%, and 79.2% respectively.

The following figure is the comparison test results of two algorithms. The driving condition is WLTP, and the SOH of the battery is 100%.

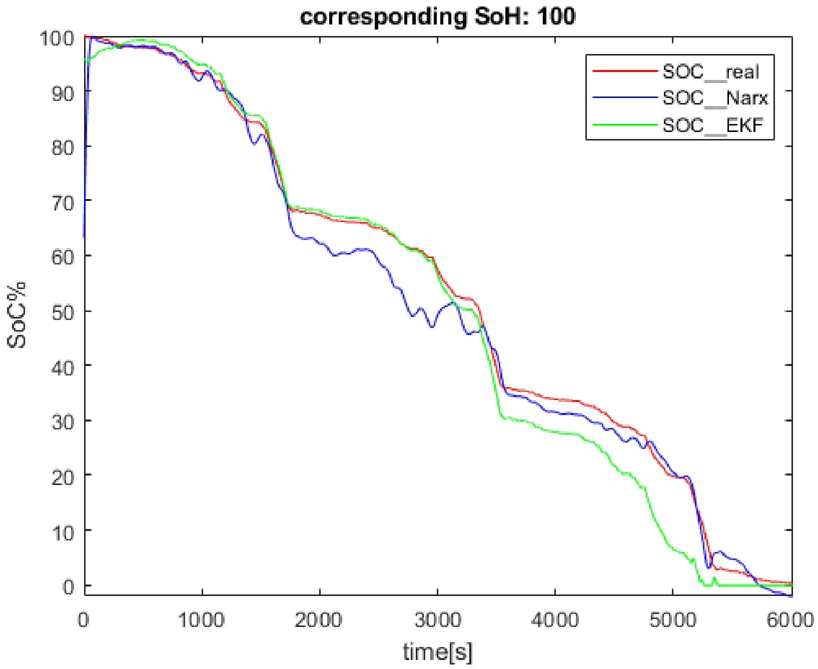


Figure 6.2.1: test results in WLTP driving condition, SoH = 100%

It can be observed from the figure that both algorithms can get reliable results from an uncertain initial value, the average error of Extended Kalman Filter is: 2.56% , the average error of Neutral Network is: 2.27%, since Neutral network can start from a completely unknown initial condition, And EKF algorithm needs the initial condition to be in the range above and below the real value, it seems that the performance of Neutral Network is better than EKF. However, in this test, the SoH of the battery is 100%, which is also a dataset for training the neural network. Therefore, to reach a conclusion, it is necessary to avoid the overlap of the test set and the training set. The following test avoids such situation:

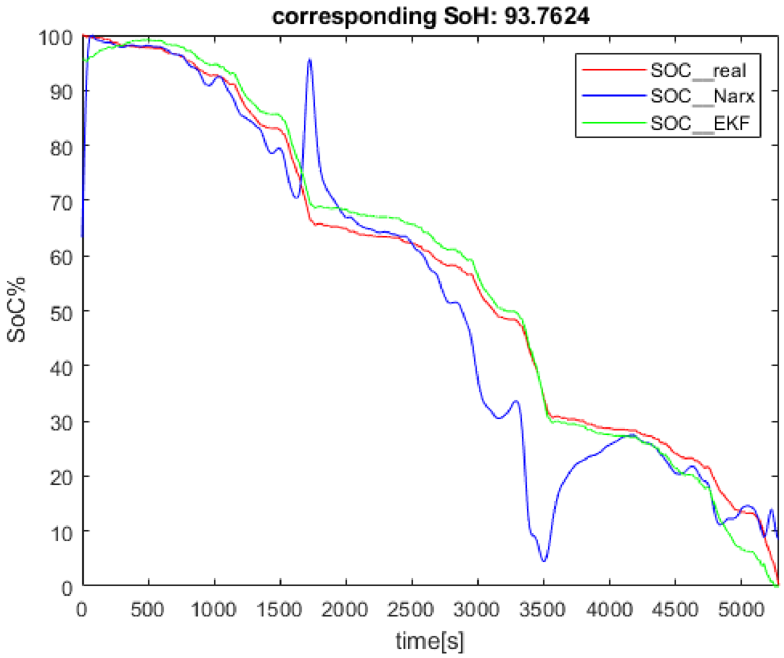


Figure 6.2.1: test results in WLTP driving condition, SoH = 93.7624%

In this case, EKF algorithm still maintains high performance, meanwhile the result from Neutral Network is poor, this is because Artificial intelligence method requires a large number of datasets to train the network, and the accuracy of the results strongly depends on the size of the test dataset. Since the dataset used to train the network is small, the results on the test set are not satisfactory.

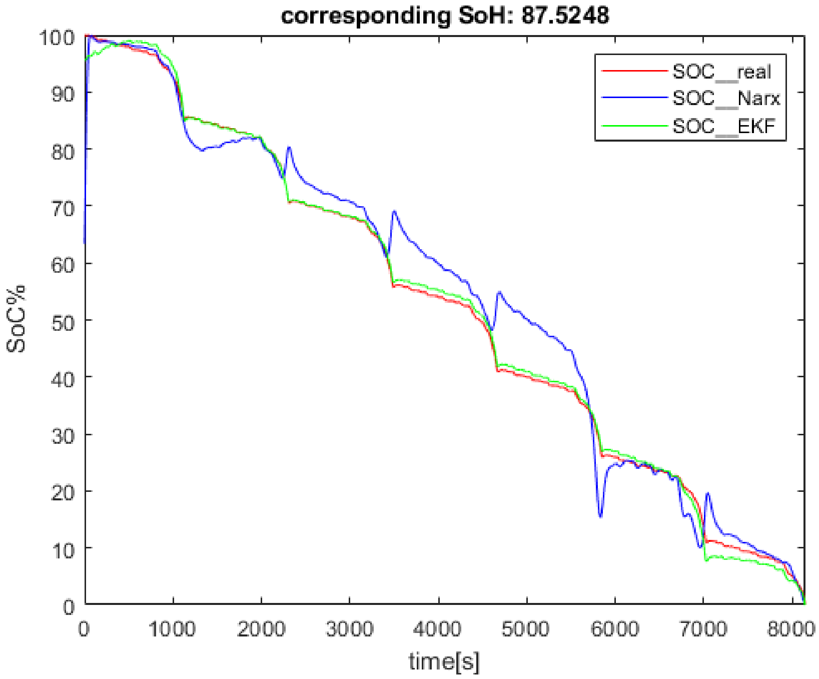


Figure 6.2.1: test results in NEDC driving condition, SoH = 87.5248%

The results obtained from the NEDC driving condition further verify the previous conclusions. Since the training dataset is completely provided by WLTP driving condition, when the test is other driving situations, the performance of the Neural Network is significantly reduced, such as the NEDC test in the figure above shows the volatility of Neutral Network result is huge, and the prediction become totally unreliable.

6.3 on road tests

The previous chapters have proved the effectiveness and robustness of the EKF algorithm. However, to apply the algorithm to practice, on-road test and Hardware in Loop test (HIL) are essential.

Figure 6.3.1: Test vehicle, a Fiat panda that has been electrified.

The vehicle we used is an older version of Fiat panda, mainly due to cost considerations, the vehicle's power system, such as engine, fuel tank, etc., are replaced with motors, batteries and other electrical components,
The vehicle has been driven in different road conditions, the battery current, terminal voltage, and battery average temperature are extracted from the CAN bus, and these data are obtained in a Battery Management System, which provided by a supplier, that has been embedded in the vehicle electrical system,

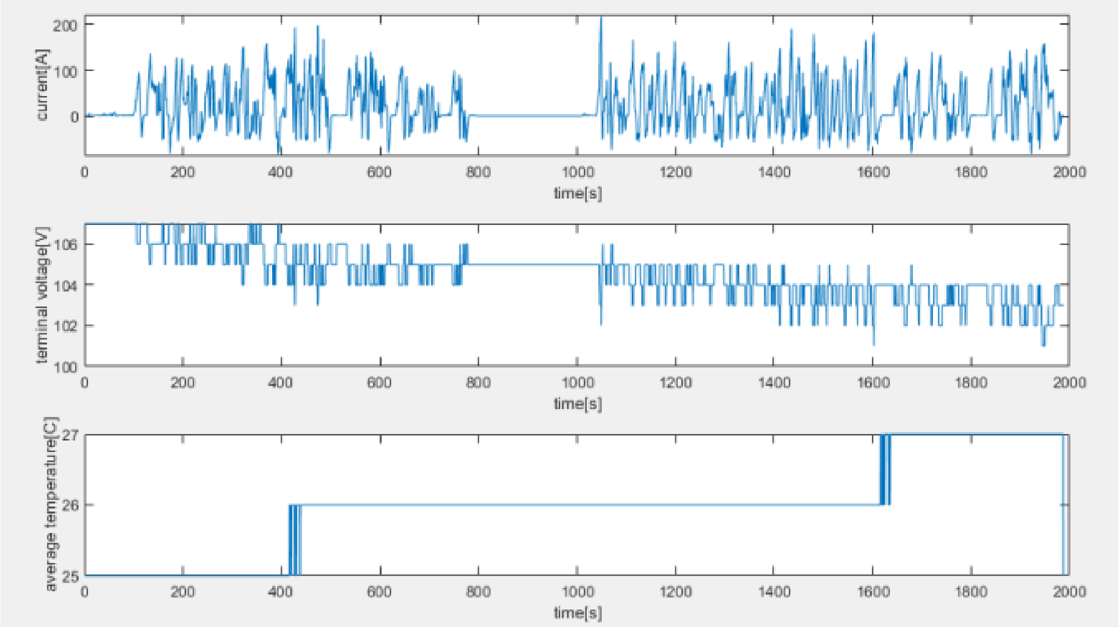


Figure 6.3.2: an example of the algorithm inputs

The target is to compare the performance of our EKF algorithm with algorithms in existing BMS. and use coulomb counting method as reference SOC, the following figures are the predicted SoC obtained by different estimate algorithms in different road conditions.

Test 1: Urban

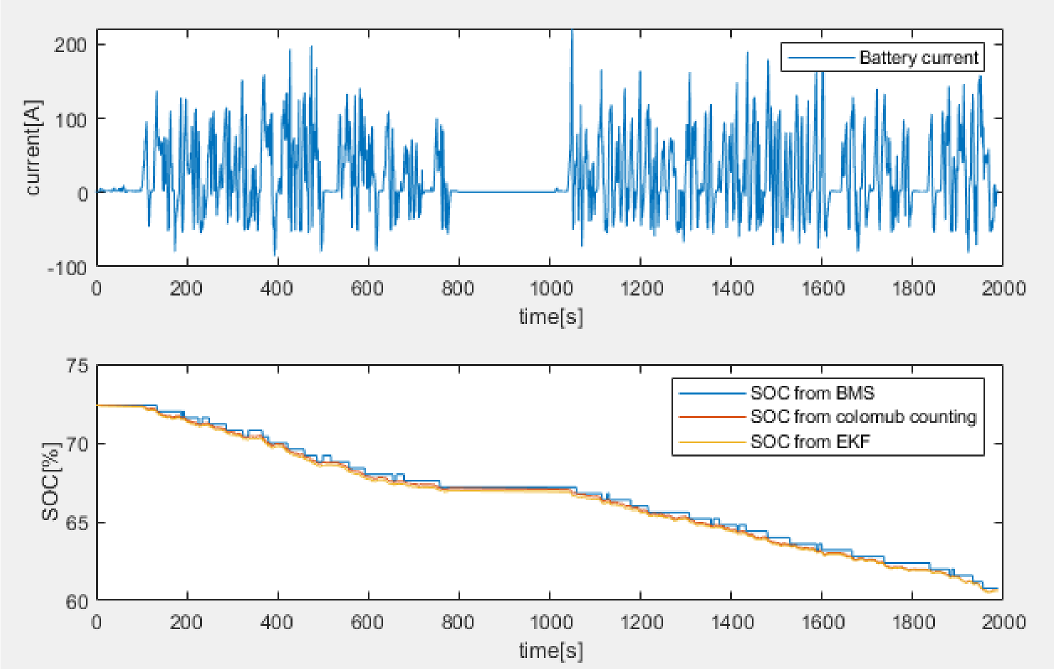


Figure 6.3.3: SoC prediction results in Urban condition

Test2: Maddalena uphill

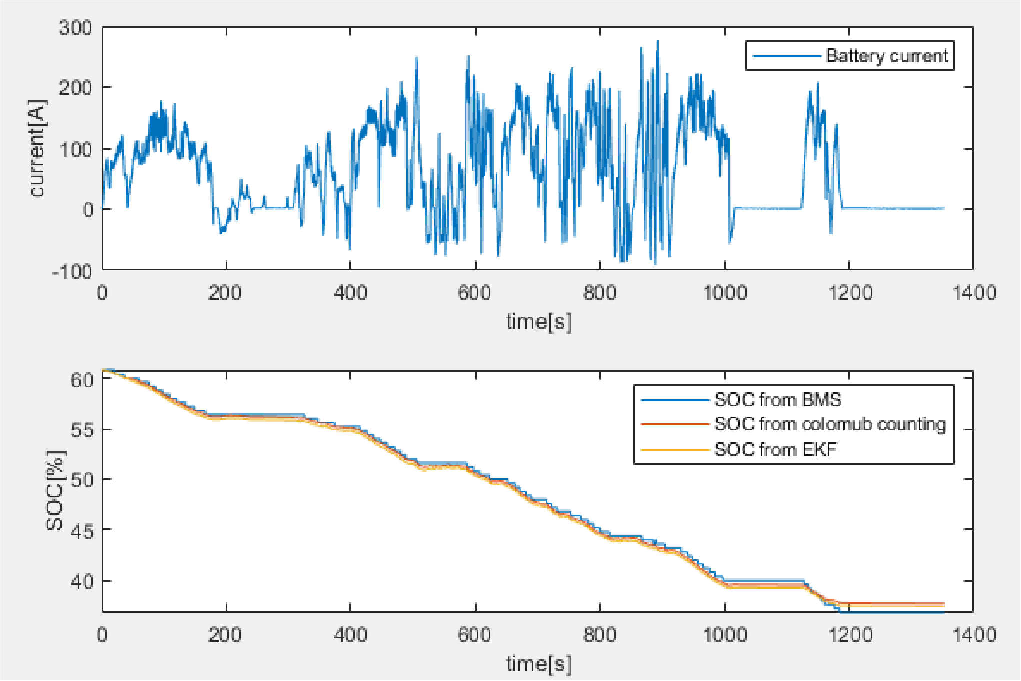


Figure 6.3.4: SoC prediction results in Maddalena uphill condition

Test3: suburban

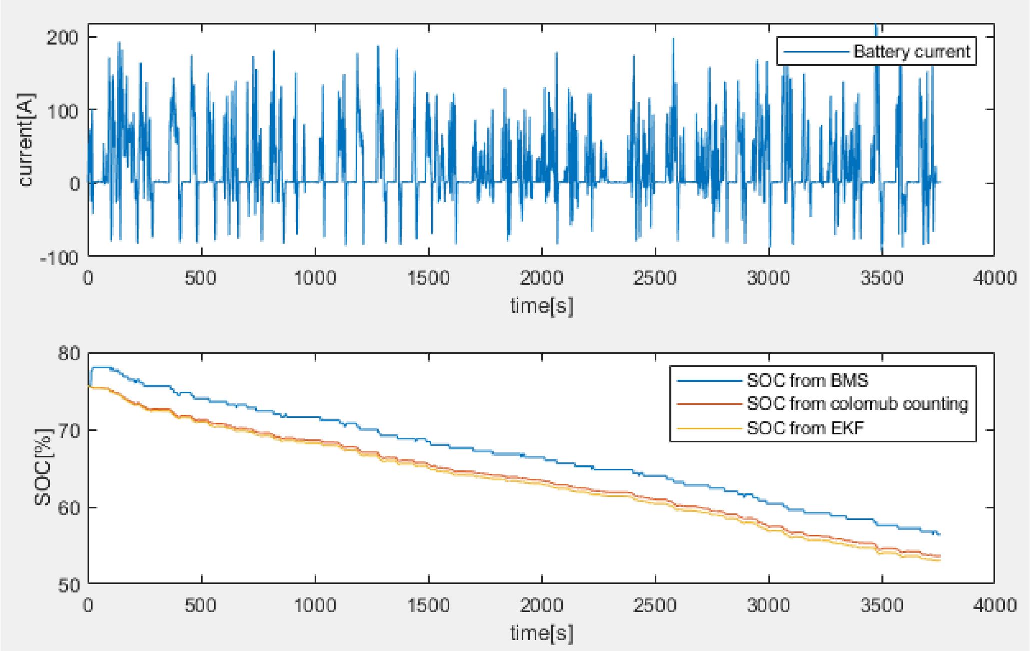


Figure 6.3.5: SoC prediction results in suburban condition

Test4: highway

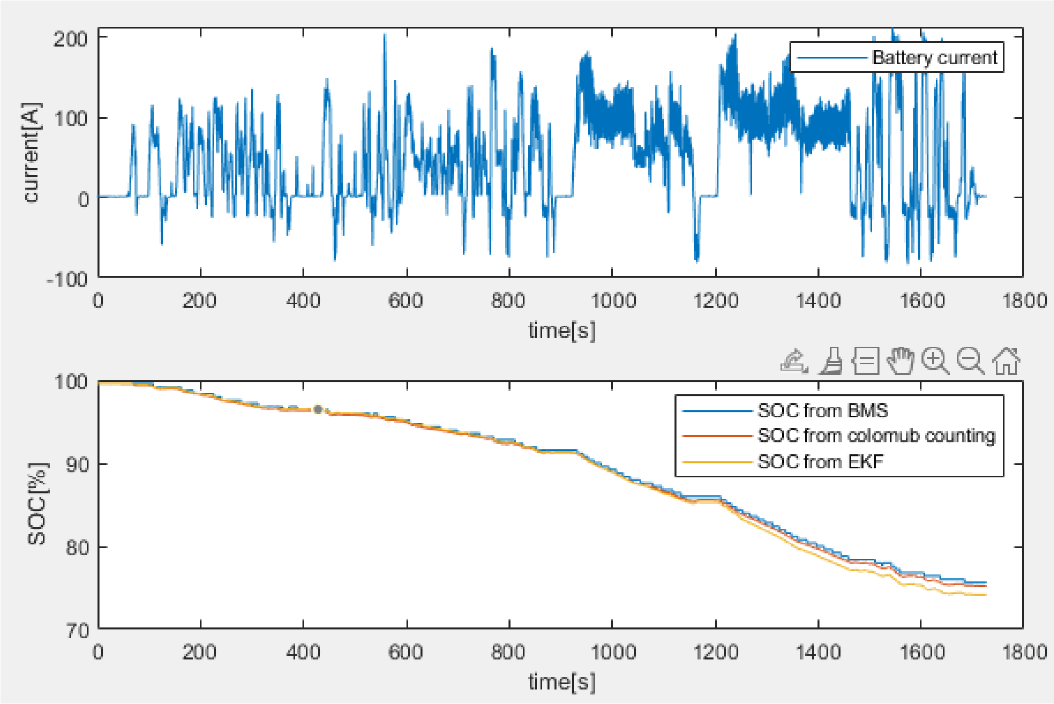


Figure 6.3.6: SoC prediction results in highway condition

Test5: urban-suburban flat

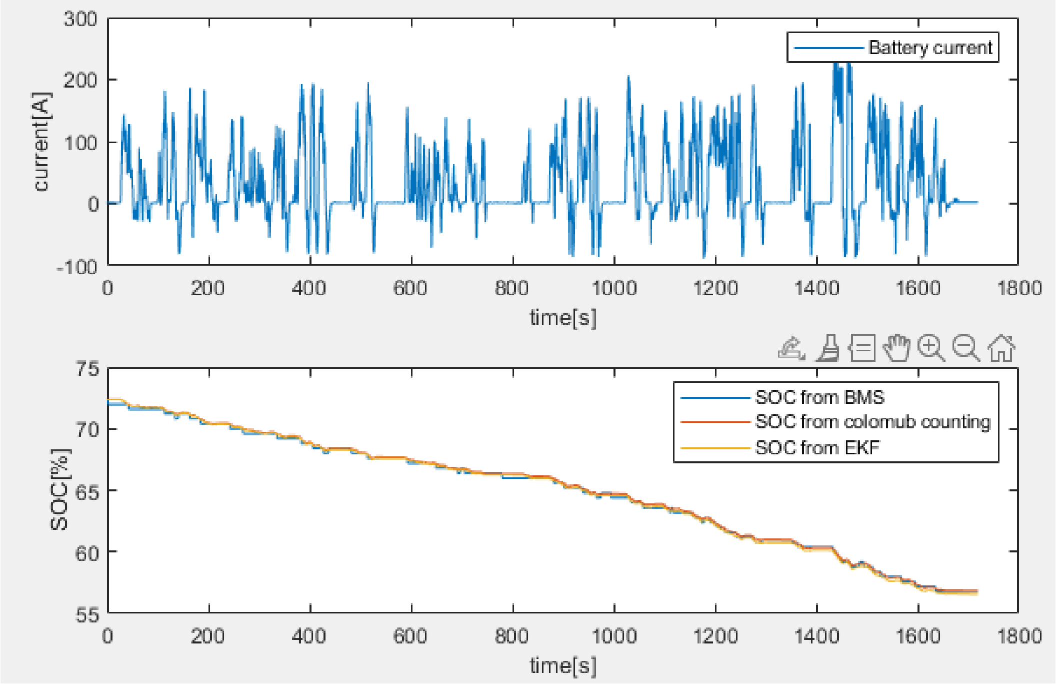


Figure 6.3.7: SoC prediction results in urban-suburban flat condition

The above figures show the situation of the vehicle driving on different driving conditions, respectively: urban, uphill, suburban, highway, urban-suburban, it can be seen that the result of our extended Kalman filter algorithm is always tracking coulomb counting , which we take as ground truth method, so this verifies that EKF is an effective algorithm in estimate SoC. Even in the suburban case, the BMS result has some deviation from coulomb counting. At the beginning of the test, the BMS result shows a significant rise in SoC, but there is no obvious charge when observing the battery current.

6.4 tests on test bench

To further verify the performance of the algorithm, a test on the test bench has been proposed.

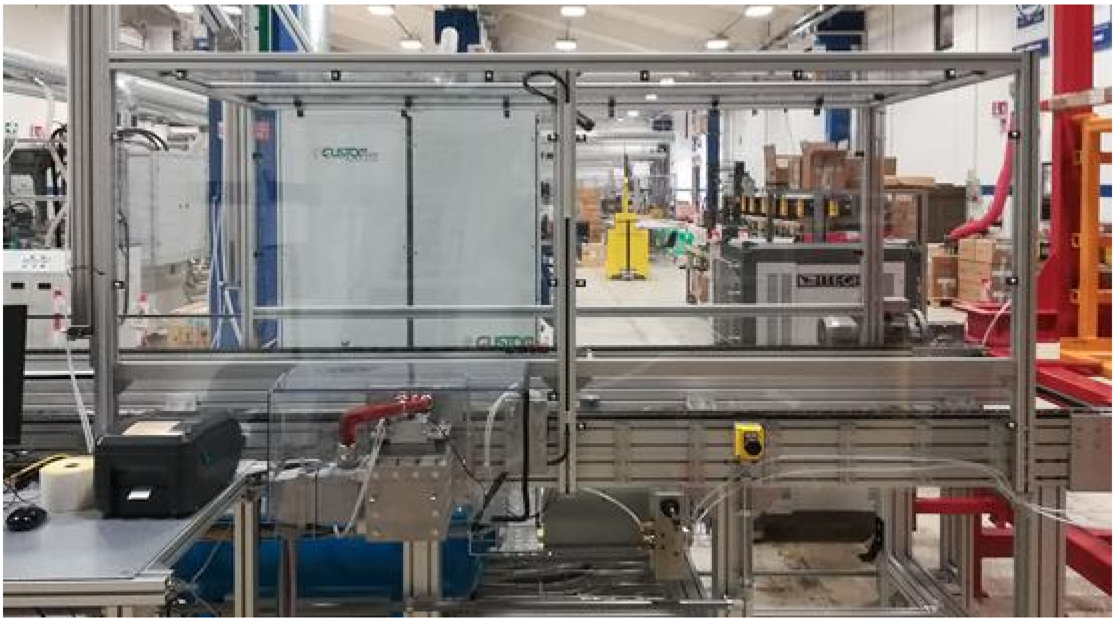


Figure 6.4.1: battery test bench

Testing follows the steps below:
1. Fully discharge the battery then fully charge it in order to get the initial SoC, the capacity of the battery in fully charged case should be compared with the maximum capacity of the battery in order to get the reference SoH
2. Discharge the battery by using WLTP class 3 driving cycle, during discharging using coulomb counting method to get the reference SoC
3. during discharge, the measured current, voltage, temperature on the battery pack level will be the input of the proposed algorithm, the output of the algorithm, SoC and SoH will be compared with their reference value to validate the performance.

Next figure shows the fully charged cycle:

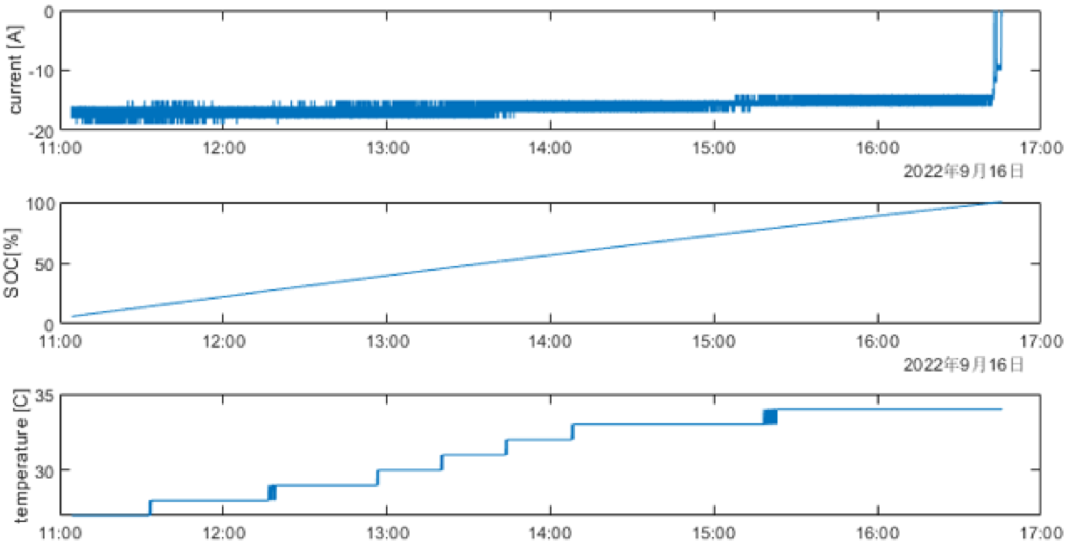


Figure 6.4.1: The charging current and the corresponding SoC

After discharging the battery , it was charged with 1.8 KW Ac charger (17 A) up to 100 % SoC, , Since the battery needs a threshold SoC to allow the control logic, it was charged from 6% remain SoC.
After integral the current we can get 92.16 [Ah] corresponding 94% SoC, In this way, the corresponding capacity of the battery at 100% SoC can be simply calculated as **98.04 [Ah],** and the reference SoH is **98%**

After getting the maximum capacity and the SoH of the battery, the next step is to simulate the discharge of the battery under a certain driving situation, at this point, a set of current profiles have been extracted from vehicle virtual model for WLTP cycle. Discharging current has been limited to 155 A since DC power supply has a maximum capacity of 18 Kw. In addition, charging current has been limited to 79A since battery charge current limit is 80A (5 sec).

The following figure shows the result of the Extended Kalman Filter algorithm predicting the SoC. Taking coulomb counting method as a reference, it can be seen that the predicted value of EKF always tracks the reference value, and the algorithm is effective.

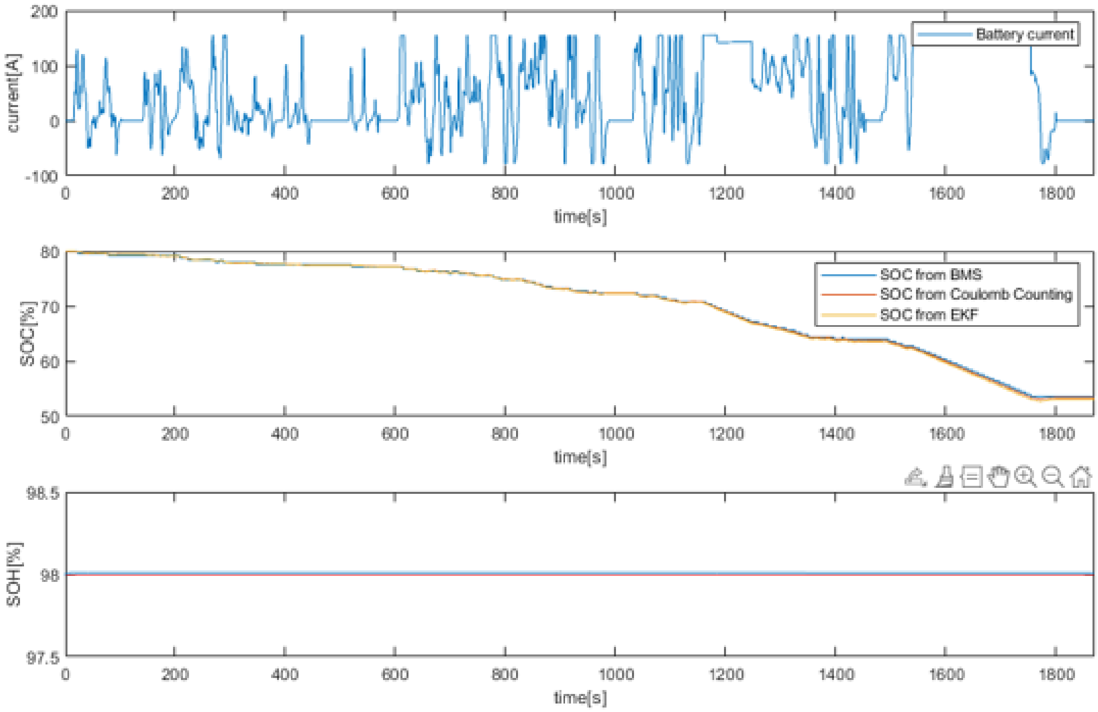
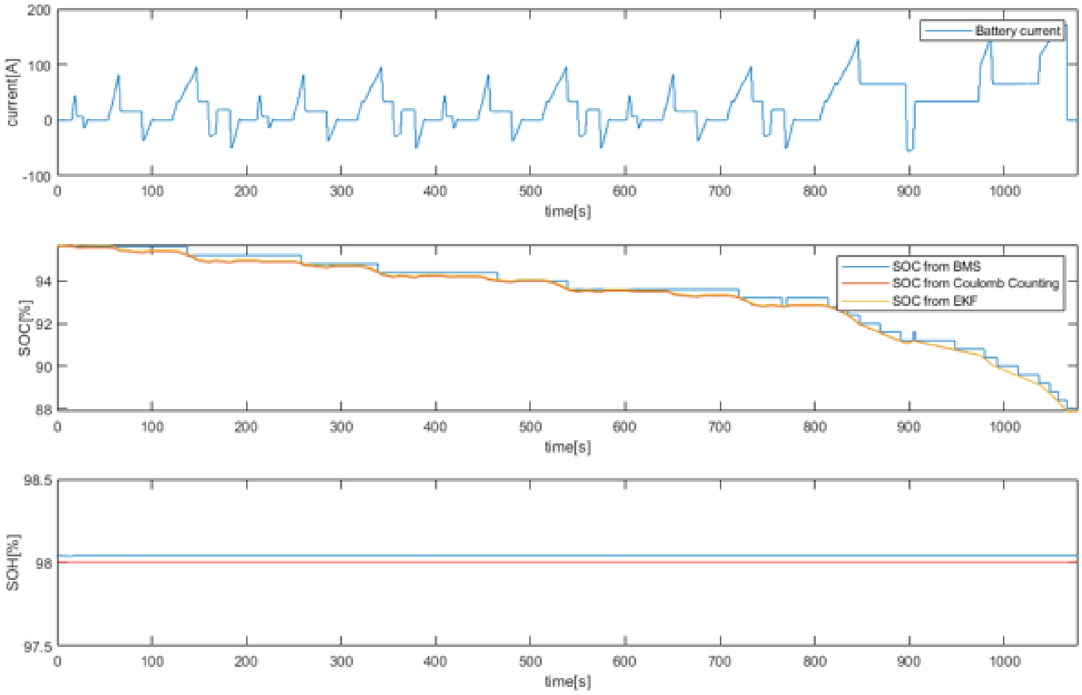


Figure 6.4.2: EKF predicting results under WLTP discharging cycle

The same discharging test using NEDC road profile to further verify the algorithm:



. Figure 6.4.3: EKF predicting results under NEDC discharging cycle

It can be seen from the test bench tests that our EKF algorithm is effective and accurate for SoC estimation in both WLTP and NEDC driving condition. For SoH, it takes a longer time to change, so it is difficult to verify its accuracy, it is recommended to use at least one full discharging cycle to verify

Chapter 7

conclusion

In this thesis work a proposed Extended Kalman Filter to estimate the State of Charge and State of Health of the battery has been developed and validate, the battery has been chosen Samsung SDI 28S2P, via several constant discharge tests and pulse discharge tests to identify the parameter of equivalent circuit model, which is a first order thevien model.

According to several standard driving cycle (WLTP,NEDC,FTP45…) dynamometric tests performed in virtual environment, the Extended Kalman Filter algorithm can reach a good precision in estimate the states of the battery, in which the average error of SoC is limited in 5%, meanwhile the average error of SoH is less than 3%, the performance is good enough for application.

Then some tests based on the experimental data are performed in order to further verify the performance and robustness of the Extended Kalman Filter, which include the on-road tests and the tests on test-bench, by comparing with an exist battery management system with an unknown algorithm for battery states estimation.

The results shows that both the EKF algorithm and the BMS algorithm are effectiveness solution for estimate the states of the battery, however the biggest advantage by using Extended Kalman Filter is that can avoid the noise from the sensor.

A possible proposal for a future work can be to test the proposed techniques on -line, if the algorithm can perform well also on the on-line tests , it means it would be an algorithm can apply in commercial.

References:

1. Battery state of health and state of charge estimation: comparison between classical and machine learning techniques, Domenico calucci
2. https://www.quattroruote.it/
3. BYD CE32BNCD lithium-ion battery cell specification report
4. Lecture E-powertrain components, prof Ravello

5. https://www.sony.com/en/SonyInfo/CorporateInfo/History/SonyHistory/2-13.html
6. Improved Performance of Li-ion Polymer Batteries Through Improved Pulse Charging Algorithm- Judy M. Amanor-Boadu 1 and Anthony Guiseppi-Elie
7. https://dieselnet.com/standards/cycles/wltp.php
8. Battery Model Parameter Estimation Using a Layered Technique: An Example Using a Lithium Iron Phosphate Cell-Tarun Huria
9. A Cell Level Model for Battery Simulation-Suguna Thanagasundram
10. https://zhuanlan.zhihu.com/p/48008124
11. https://zhuanlan.zhihu.com/p/48876718
12. https://zhuanlan.zhihu.com/p/67138271
13. https://zhuanlan.zhihu.com/p/550160197
14. https://blog.csdn.net/weixin_42905141/article/details/99710297
15. https://zhuanlan.zhihu.com/p/41767489
16. https://www.guyuehome.com/37396
17. Data analysis and Artificial Intelligence- TATIANA TOMMASI
18. Comparison of One-Pedal Driving Strategies for Electric Vehicles AlirezaMoayyedi
19. Samsung 50Ah cell datasheet
20. understanding and applying coulombic efficiency in lithium metal batteries – Jie xiao, natural energy
21. https://ww2.mathworks.cn/matlabcentral/fileexchange/34765-polyfitn