UPPSALA
UNIVERSITET

UPTEC F 21022
Examensarbete 30 hp
Juni 2021

Linear parameter varying model and identification method for Li-ion batteries in electric vehicles

Per Larsson

Civilingenjörsprogrammet i teknisk fysik

UPPSALA UNIVERSITET

Linear parameter varying model and identification method for Li-ion batteries in electric vehicles

Per Larsson

Abstract

The market for electric vehicles (EV) is growing rapidly. The rise of EVs is most prominent for passenger vehicles, but trucks and busses are also quickly becoming electrified. Scania aims to be the leader of this transition. A central part of the EV is the Lithium-ion battery. In order to use the battery in the most efficient manner a Battery Management System (BMS) is needed. A key part of the BMS is a model that describes the battery as a system where the input is the current and the output is the terminal voltage. The dynamics of the battery is affected by external factors, called scheduling variables, that should be taken in to account in order to acquire an accurate model. This thesis aims to capture this behavior by the identification of a Linear Parameter Varying (LPV) model that has State of Charge (SOC) and temperature as scheduling variables. The LPV model was identified by first performing a set of local system identifications at varying levels of the scheduling variables. From this, a set of different LPV model structures were set up and then optimized with the use of datasets with a wider coverage of the scheduling variables. The results showed that there are clear advantages in using an LPV model compared to a traditional constant model, but that the robustness of the model largely is dependent on the choice of the data used for optimization.

Teknisk-naturvetenskapliga fakulteten

Uppsala universitet, Utgivningsort Uppsala/Visby

Handledare: Nikolaos Karavalakis Ämnesgranskare: Per Mattsson

Examinator: Tomas Nyberg

Populärvetenskaplig sammanfattning

Marknaden för elektriska fordon växer snabbt. De elektriska fordonens framväxt är tydligast när det kommer till personbilar, men lastbilar och bussar blir i en allt större utsträckning elektrifierade. Scania har som målsättning att vara ledande i denna omställning. En central del av det elektriska fordonet är Litiumjonbatteriet. För att använda batteriet på ett så effektivt sätt som möjligt krävs ett styrsystem som övervakar batteriet. En viktig komponent i detta styrsystem är en model av batteriet som beskriver batteriet som ett system med strömmen som dras från batteriet som insignal, och polspänningen som utsignal. Batteriets beteende påverkas av externa faktorer, och denna påverkan bör tas i beaktning för att kunna skapa en bra modell. Detta projekt ämnar att beskriva detta beteende genom att identifiera en så kallad *linear parameter varying*-modell som har batteriets laddningsgrad och temperatur som de externa faktorerna. Tillvägagångssättet för att identifiera denna modell var att först identifiera ett antal lokala modeller vid olika värden på laddningsgraden och temperaturen. Från detta kunde sedan ett antal olika modellkandidater tas fram. Dessa modeller optimerades med hjälp av data med stor variation i laddningsgraden och temperaturen. Resultaten visade att det finns klara fördelar med att använda denna typ av modell i jämförelse med traditionella konstanta modeller. Dock visade resultaten också att stor vikt bör läggas vid valet av datan för optimering för att få en robust modell.

Acknowledgements

Writing a master's thesis during a global pandemic has been challenging, meaning the support I got meant even more.

I would like to thank my supervisor Nikolaos Karavalakis for his devoted support of my work. Nikolaos trusted me to make many choices in the model design process, but also gave me valuable guidance when needed.

I would also like to thank my subject reader Per Mattsson for his valuable input on the more theoretical aspects of the project. Per helped me see the project from new angles.

Lastly I would like to thank all the other Scania employees that has aided me during my thesis. Special thanks to Malin Andersson that shared with me her previous experience of modelling of linear parameter varying systems.

Table of contents

List of Figures

List of Tables

List of Abbreviations

EV	Electric Vehicle
BMS	Battery Management System
SOC	State Of Charge
SOH	State Of Health
SISO	Single-Input-Single-Output
MIMO	Multiple-Input-Multiple-Output
SS	State-Space
LTI	Linear Time Invariant
NL	Non-Linear
LPV	Linear Parameter Varying
LTV	Linear Time Varying
OCV	Open Circuit Voltage
ECM	Equivalent Circuit Model
RPT	Reference Performance Test
HPPC	High Power Pulse Characterization
NRMSE	Normalized Root Mean Square Error
DC	Drive Cycle

Chapter 1

Introduction

The global EV Outlook is an annual publication that identifies and discusses recent developments in electric mobility around the world [1]. In their 2020 publication they concluded that the market for electric vehicles (EV) keeps growing rapidly. The growth is largely driven by policies and announcements from governments where the zero emission strategies motivates subsidies for investments in electrification of the transportation industry. This is the reason why the infrastructure for electric vehicle charging continues to expand. Even if the rise of EVs are most prominent for passenger cars, trucks and busses are also quickly becoming more electric. Scania aims to be the leader of the transition in to electric driven buses and trucks, with many options already available for hybrid electric or fully electric propulsion units for its vehicles. Scania is also exploring alternative solutions for electrification, like the continuous charging along electric roads, and fuel cell trucks [2].

1.1 The Lithium-ion battery

Lithium-ion (Li-ion) batteries has an unmatchable combination of both high power and energy density compared to other types of batteries. The reason for this lies in the chemical properties of lithium, which has a high reduction capacity and is also light. This is the reason why Li-ion batteries are the most attractive choice for use in portable electronics, power tools and electric vehicles. There is an immense interest in the usage and development of Li-ion batteries both from the industry and government funding agencies since the usage of them could reduce greenhouse gas emissions. There is also a possibility of using the batteries as a power backup in the electric grid. The main obstacle in the increased use of lithium in batteries is the relatively high cost of the metal. While there is doubt whether the supply of lithium is sufficient, it is argued in [3] that there still exists enough lithium in the earth's crust to power a global fleet of electric automobiles. Since Li-ion evidently is the best currently known choice for batteries, the effort to optimize their performance and efficiency is of high priority.

In order to be able to use the battery in a safe and efficient manner a battery management system (BMS) is needed. The purpose of the BMS is to oversee the operation of the battery by monitoring the voltage of each cell in the battery pack, the current drawn from the battery and the temperature. The BMS also monitors abusive behavior that may damage the battery. Using all this information the BMS ensures that the battery is

not overcharged or undercharged by continuously estimating the State of Charge (SOC) of the battery. The measurements are also used to estimate the battery's state of health (SOH), which is a measure of the capacity and efficiency that gradually decreases as the battery is used. In order to be able to compute all these relevant statistics the BMS must be based on an accurate model of the battery that represents how the cells behave when different currents are applied, cf. [4].

The battery cell can be viewed as a Single-Input-Single-Output (SISO) system where the applied current is the input, and the terminal voltage of the cell is the output. In isothermal conditions with constant SOC and SOH this system is well described by a linear time invariant (LTI) system. This is however mostly not the case since temperature is affected by the current drawn from the battery, as well as the ambient temperature around the cell. As SOC is a direct function of the history of the current applied to the battery, it also changes during operation. This introduces the need for a model of the battery that captures the battery's non linear behavior with respect to these parameters.

1.2 The linear parameter varying model

Many systems that need to be monitored or controlled have a non-linear (NL) behavior. While there exist NL models, they do not have the advantage of robust and optimal control that LTI models offer. This makes LTI models favorable to NL models in most cases. One way to use the LTI structure but capturing the behavior of a NL system is to use a linear parameter varying (LPV) model. The LPV model is based on a LTI model, but has the property that the model behavior is decided by some function of one or several external parameters. These parameters are called scheduling variables. This means that the LPV model is a LTI model as long as the scheduling variables are fixed, but continuously changes its behavior as the scheduling variables change. If the scheduling variables depend on the state or output variables the system is known as quasi-LPV [5]. The possible values of the scheduling variables spans what is called the scheduling space. This concept is illustrated in Figure 1.2.1. In the figure the blue sheet represents the scheduling space, the red arrows are the scheduling functions and the green boxes are local LTI models. The scheduling space in the figure has specifically two dimensions p_1 and p_2 because this aids the illustration, but the scheduling space can have any natural number of dimensions. By interpolating between these local models the scheduling functions can be built up. From the scheduling space the mapping to the different models are then done via the defined scheduling variable dependency. One way to describe the LPV model is that it is an extension of linear time varying (LTV) models [5]. The time varying aspect is then the change in model dynamics caused by the scheduling variables. The identification of LPV models lacks a general theory. This is largely due to the fact that there does not exist a structured framework for the description of the model class. LPV modelling also requires several assumptions when it comes to the dependency on the scheduling variables. Despite this, LPV models have been successfully implemented in the industry [7].

Identification of LPV models require, like for the case with a regular LTI model, datasets with input and output data. By employing some sort of optimization procedure the model is then fitted to the dataset. Alongside this, the scheduling variables need

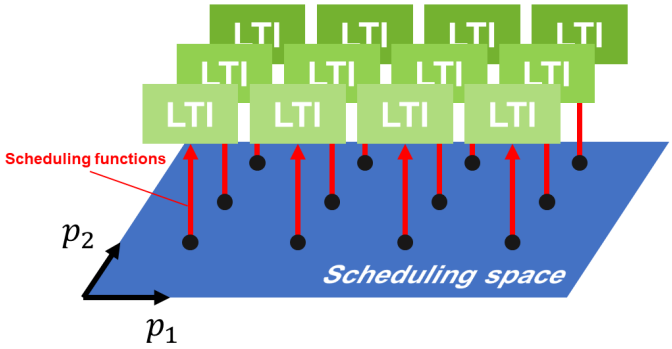


Figure 1.2.1: Illustration of the mapping from the scheduling space to the model space.

to be known and part of the dataset. System identification of LPV models is divided into two approaches, the local approach and the global approach. The local approach aims to identify a set of models with the scheduling variables kept constant for each individual identification procedure. The scheduling variables are then altered for the next identification. By varying the scheduling variables so that the whole scheduling space is spanned, one then has a set of local models. From this set one can then interpolate between the models and acquire the LPV model. The global approach instead aims to identify the complete model in one identification procedure by continuously varying the scheduling variables throughout the dataset used for identification. This subsequently involves less effort for the identification of the LPV model but does require that the scheduling variable dependence is known beforehand [6].

1.3 Thesis objective

The objective of this thesis is to identify a LPV model for the Li-ion battery. The number of model parameters should be as small as possible, as this is key in keeping the model simple for identification and reducing the memory footprint. The number of model parameters should nevertheless be big enough to achieve a good level of accuracy. The thesis should also clearly describe the identification procedure to lay a good foundation for further investigation of LPV modelling at Scania. LPV models have previously been used to describe the Li-ion battery [8–14] with good results. These studies use different configurations of scheduling variables, but all conclude that the use of the LPV model is beneficial, which suggests that a LPV model could be used as a Li-ion battery model for the BMS used in Scania's EVs.

Chapter 2

Theory

In this chapter the relevant system identification and battery modelling theory will be provided along with the methods for parameter extraction on the battery. The parameter dependence found in earlier studies will be described along with LPV models previously used for the Li-ion battery. Lastly, the identification techniques used in previous studies with LPV-modelling of the Li-ion will be described.

2.1 System identification

This section will explain some fundamental steps and properties of the system identification process used in this thesis.

2.1.1 Removal of open circuit voltage bias

The voltage of the battery cell can be divided into two parts, one static and one dynamic. The static part is called the open circuit voltage (OCV). The OCV is a direct function of the SOC. The dynamic part is called polarization voltage and is characterized as any deviation from the OCV. After long periods of rest, i.e. no current is applied to the battery, the battery voltage converges to the OCV. One key point with modelling the battery cell is to be able to subtract the dynamic part of the voltage from the terminal voltage in order to acquire the OCV. Knowing the relationship between the OCV and SOC can then allow for the estimation of the cell SOC, which is necessary for the battery usage in an EV. Because of this, the model to be designed should model the dynamic part of the voltage.

SOC is naturally affected by the energy throughput, which charges or discharges the battery depending on current direction. An ideal SOC estimator would just integrate this current over time to find the SOC by

$$z(t) = z(t_0) + \frac{1}{Q}\int_{t_0}^{t} \eta(\tau) i(\tau) d\tau, \tag{2.1.1}$$

where $z(t)$ is the SOC at time t, t_0 is the starting time of the current measurement, Q is the capacity of the battery cell, $\eta(\tau)$ is the coulombic efficiency at time τ and $i(\tau)$ is the

current applied to the cell at time τ. Charging current is defined as positive [15]. This integration procedure is called coulomb counting and can be trusted to be sufficiently accurate under laboratory conditions. During use in an EV, this process is however not trustworthy because of measurement noise and biases that shifts the SOC estimation away from the true value.

2.1.2 Goodness of fit

A common way to measure model performance is the Normalized Root Mean Square Error (NRMSE) measure defined by

$$NRMSE = \frac{||y - \hat{y}||}{||y - \bar{y}||}, \tag{2.1.2}$$

where y is the measured output of the system, $\hat{y}$ is the simulated output and $\bar{y}$ is the mean of the output. By computing

$$fit = 100(1 - NRMSE), \tag{2.1.3}$$

one gets an indication of how well the simulated output corresponds to the measured output, where 100% is a perfect fit and 0% means the simulated output is as close to the measured output as the mean of the output. The fit consequently gives negative values if the simulated output is worse than the mean of the measured output.

2.1.3 Linear model and linearity in parameters

The LPV model is linear in the system dynamics for fixed scheduling variables, as the name suggests. This should not be confused with how the scheduling functions are designed. The scheduling functions are just some set of functions that map one point in the scheduling space to a LTI model. While this mapping can be linear, it does not have to be. Another relevant distinction to make is whether the model is linear in the parameters or not. For the model to be linear in the parameters, the elements of the system matrices must be arranged so that they can be written as a linear regression. More specifically this means that if the scheduling functions directly decide the system matrices' elements, then the scheduling functions needs to be linear in the parameters.

A key aspect with having the model be linear in the parameters is that the model is more easily identified with e.g. a least-squares algorithm, due to the possibility of writing the system matrices as a linear regression.

2.1.4 Gray-box model

A white-box model is a model that is based on prior knowledge about a given system's dynamics. This means that no data has to be acquired from the system in order to describe its input-output relationship. By means of previously established physical relationships and laws, the model dynamics can be specified. A black-box model on the other hand is completely based on measured input and output data and is essentially a data description model. A black-box model is only designed to describe the measured output as closely as possible. The gray-box model, as the name suggests, incorporates both knowledge of the system dynamics as well as measured input and output data [16].

2.2 The equivalent circuit model

One common way to model the battery is to use an equivalent circuit model (ECM), which is a gray-box model. The ECM is not a representation of what actually exist inside the battery, but a tool for describing its behavior. The ECM is widely used because of its simplicity and yet good ability to capture the battery dynamics [8, 10, 11]. The ECM can be constructed in a variety of ways. This is a trade-off between complexity and accuracy. In Figure 2.2.1 a circuit diagram of the ECM of order n can be seen. The model incorporates an ideal voltage source that represents the OCV. Connected in series is the cells ohmic resistance R_0, and a number of RC networks with resistance R_i and capacitance C_i. The terminal cell voltage of the battery, i.e. the voltage that is measured is denoted V_t. From this model it is possible to identify the voltage drop V_0 over R_0 to be

$$V_0 = R_0 I, \tag{2.2.1}$$

where I is the current flowing through the resistor. Further on the first order differential equation that describes the voltage drop V_i over RC network i is given by

$$\frac{dV_i}{dt} = -\frac{1}{R_i C_i} V_i + \frac{1}{C_i} I, \tag{2.2.2}$$

where R_i and C_i is the resistance and capacitance of RC network i [8]. As the dynamic part of the voltage is of interest, the OCV can be removed from the terminal voltage by

$$V_t - OCV = V_0 + \sum_{i=1}^{n} V_i. \tag{2.2.3}$$

From this the state space model

$$\dot{x} = \begin{bmatrix} -\frac{1}{R_1 C_1} & 0 & \dots & 0 \\ 0 & -\frac{1}{R_2 C_2} & & \\ \vdots & & \ddots & \\ 0 & & & -\frac{1}{R_n C_n} \end{bmatrix} x + \begin{bmatrix} \frac{1}{C_1} \\ \frac{1}{C_2} \\ \vdots \\ \frac{1}{C_n} \end{bmatrix} u \tag{2.2.4}$$

$$y = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} x + R_0 u, \tag{2.2.5}$$

can be constructed, where the state vector, input and output are

$$x = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} \qquad u = I \qquad y = V_0 + \sum_{i=1}^{n} V_i. \tag{2.2.6}$$

The ohmic resistance models the instant change in voltage that the battery shows when current is applied. The rest of the change in voltage is caused by polarization. The RC networks model the diffusion voltages that are part of the polarization that occurs in the cell when it is subjected to a load. The number of networks is largely a factor of the accuracy of the model [15]. The other part of the polarization is the so called hysteresis. When letting the battery rest after use, making the diffusion voltages fade

out, the terminal voltage will not quite converge to the OCV. Depending on if the battery was charged or discharged before the resting period, the voltage will either be slightly above or below the OCV. This is called hysteresis [8, 15]. While the diffusion voltage will decay with time, the hysteresis is constant even after resting of the battery cell. This makes modelling of the hysteresis complicated, and it is thus excluded from the model described here.

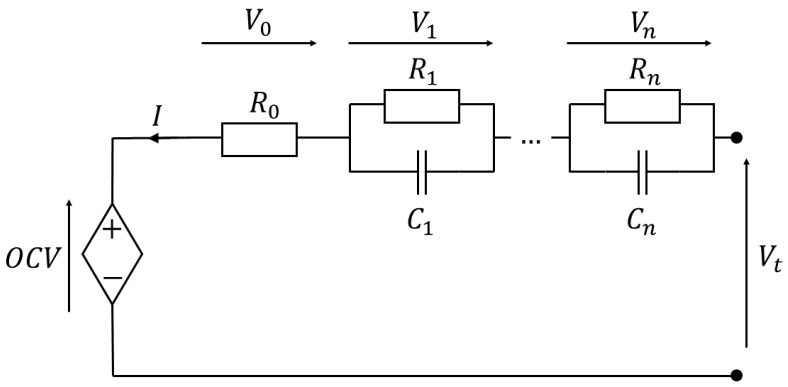


Figure 2.2.1: The equivalent circuit model of order n, and thus having n RC networks.

2.3 Parameter extraction

An analytical way of identifying the parameters of the ECM is to analyze reference performance tests (RPT). The RPTs are carefully carried out in a lab environment where the current, voltage and temperature signals can be measured accurately. This means that the measurement noise is low, and the data acquired from the test is well suited for system identification purposes. A common way to structure the RPT is the hybrid power pulse characterization (HPPC) [17]. In these tests constant current pulses are applied to the battery, and the voltage response is studied. By having the current input in the form of a rectangular pulse the voltage output can be studied as a step response. System identification theory then lets one analyse the step response in order to identify the circuit parameters [15, 18]. In Figure 2.3.1 a typical current pulse and voltage response in a HPPC can be seen. The ohmic resistance R_0 can be identified by observing the instant change ΔV_0 that occurs in output voltage when the current is applied. Knowing this, as well as the pulse current ΔI lets us calculate the resistance by

$$R_0 = \frac{\Delta V_0}{\Delta I}. \tag{2.3.1}$$

The shape of the the voltage response following the instant change ΔV_{diff} is characterized by the diffusion voltage modeled by the RC networks. This is a transient that starts just after the instant voltage change and lasts until a steady state voltage has been reached. Depending on how many RC networks are used the transient is divided into different

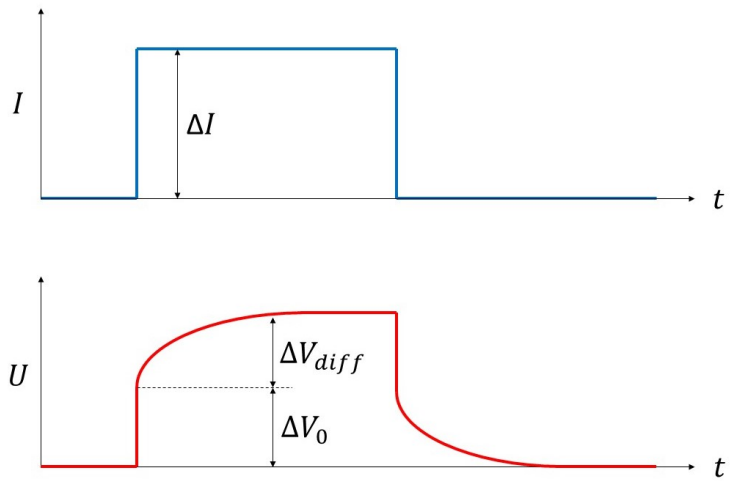


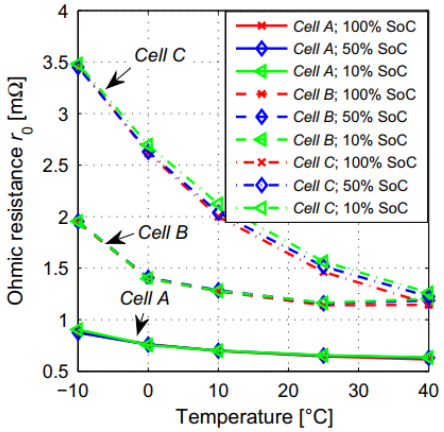
Figure 2.3.1: Typical shape of the current pulse and voltage response part of a HPPC test

regions of fast and slow transients. The first RC network models the fast part of the transient, and the higher order networks model the slower transients. Because of this the first short part is dedicated to the first order RC network, and the following time is modeled as caused by the higher order networks. By dividing the curve this way the effect of the fast and slow transient curves on the other is minimized. This division of the transient is not well defined. The study in [18] divided the transient so that the slow transients lasts four times longer than the fast transient, which gave satisfying results. From this division the time constant $\tau_i = R_i C_i$ can be identified as the time it takes for the transient to reach 67% of its steady state voltage. This steady state voltage can, together with the current ΔI, be used to find the resistance of each RC network by $R_i = \frac{\Delta V_i}{\Delta I}$, where ΔV_i is the voltage change in transient i.

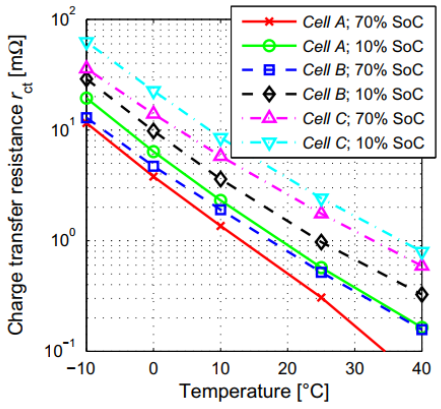
This first estimation of the circuit parameters will yield a rough estimate of the model, but for an optimal model the parameters need to be optimized given the available dataset. Examples of this optimization include both linear least squares [19] and non-linear least squares [20]. This thesis focuses on estimating the system matrices with the **ssest** function from the System Identification Toolbox in MATLAB. The estimation has to be constrained so that the C matrix is not estimated, since it is constant. The noise matrices E_1 and E_2 from 2.5.1 are also not estimated since these are not part of the model. Another necessary constraint is that the circuit parameters must be positive. Negative resistance or capacitance is not physical. Negative values on R_1 or C_1 will also lead to an unstable model, since this means that the A system matrix in 3.0.1 has a positive eigenvalue. From the local identifications the circuit parameters of the ECM can be extracted. The circuit parameter R_0 is extracted by just taking the value of the D matrix. C_1 is acquired from the inverse of the B matrix. With C_1 known, R_1 can be extracted by multiplying the A matrix with $-C_1$ and then taking the inverse.

2.4 Parameter dependence

In [21] a thorough investigation of the battery parameters' dependence on external factors is studied. It is investigated how temperature, SOC, SOH and current affects the model parameters, and more precisely how the dependence on one parameter is changed due to another. The ohmic resistance R_0 increases as the temperature decreases. This dependency on temperature is also increased by age, increasing the slope of the curve and shifting it upwards. The ohmic resistance however does not show any clear dependence on SOC for a fixed temperature. Only a first order ECM is considered in the study, so only one RC network with the parameters R_1 and C_1 is present. It is found that R_1 has a clear exponential dependence on temperature, which is explained by the exponential form of the Arrhenius equation. In contrast to the ohmic resistance it can not be said that the form of this dependence on temperature is affected by SOH, even though the temperature dependence of R_1 varies from cell to cell and changes slightly with SOH. In Figure 2.4.1 this dependence on temperature for both R_0 and R_1 can be seen. Note that in Figure 2.4.1b the y-axis is logarithmic.



(a) The dependence of ohmic resistance R_0 on temperature for three different stages of aging.

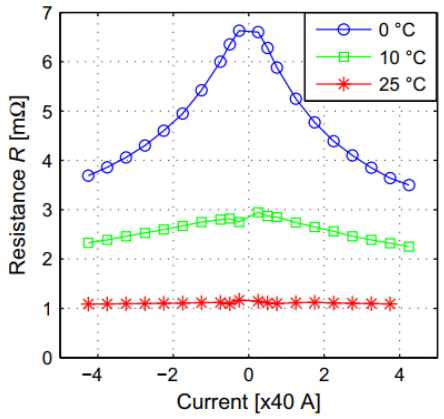


(b) The dependence of the charge transfer resistance R_{ct} (R_1) on temperature for three different stages of aging.

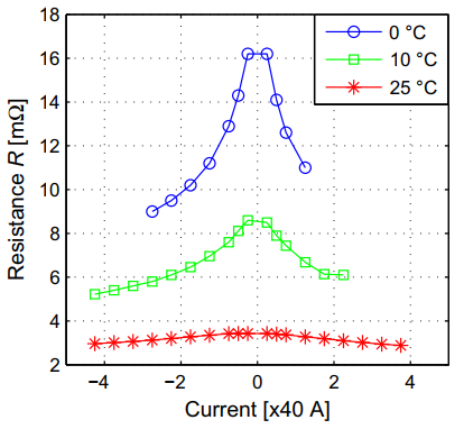
Figure 2.4.1: The temperature dependence of the ohmic resistance R_0 and the charge transfer resistance R_1 found in [21].

It can be concluded that R_1 does depend on SOC. Firstly the temperature dependence is strengthened as SOC decreases. Secondly, for a fixed temperature R_1 has a clear dependence on SOC that is greatly increased at lower SOH. For new cells the increase of R_1 is limited to only very high or low levels of SOH. For older cells on the other hand the effect of SOC is only negligible between 50% and 70% SOC. The time constant τ_1 shows an exponential dependency on temperature. This is however expected since $\tau_1 = R_1 C_1$, and the temperature dependence of R_1 is known. It is concluded that C_1 also is temperature dependent, though not as clearly as R_1. C_1 does however increase with decreased SOH. The time constant shows a similar dependence on SOC as R_1, which is backed up by the fact that C_1 does not have a significant dependency on SOC. The

parameters' dependency on the current is also investigated by applying direct current pulses to the battery and noting the voltage drop adjusted for OCV change. It is found that the total resistance of the battery, i.e. $R_0 + R_1$ does not show a clear dependence on current at room temperature for a new cell. When the temperature is decreased there does however exist a clear dependence that shows decreasing resistance as the current magnitude is increased. An older cell showed the temperature dependency even at room temperature. The dependency on current also changed under different levels of SOC, but not greatly. In Figure 2.4.2 these dependencies are illustrated.



(a) The battery resistance dependence on current for a new cell.



(b) The battery resistance dependence on current for an aged cell.

Figure 2.4.2: Battery resistance dependence on current [21].

Conclusively it can be said that all of the investigated parameters have some effect on the ECM parameters. Most prominent is the temperature affecting the resistance elements. As the temperature dependency of the internal resistance shows an exponential form at all stages of SOH it is suggested in [21] that a proportional scaling of the exponential curve would describe this relation well.

2.5 LPV model for the Li-ion battery

LPV models are most commonly represented as state-space (SS) models, since LPV control theory requires the SS representation. Using SS representation also simplifies the extension of the system to multiple-input-multiple-output (MIMO) form [5]. Consider the LPV-SS model

$$\begin{aligned} \dot{x}(t) =& A(p)x(t) + B(p)u(t) + E_1(p)e(t) \\ y(t) =& C(p)x(t) + D(p)u(t) + E_2(p)e(t), \end{aligned} \tag{2.5.1}$$

where the difference from a regular LTI-SS representation lies in the dependence on the scheduling variables p in the system matrices A, B, C, D, E_1 and E_2. It is often assumed that the noise is not dependent on the scheduling, removing the dependence on p from

E_1 and E_2. A common way to define the parameter dependence is that the relation is affine (linear). This means the system matrices in 2.5.1 take the form

$$A(p) = A_0 + \sum_{l=1}^{n_p} A_l p_l, \tag{2.5.2}$$

where n_p is the number of scheduling variables, A_l, $l = 0, 1...n_p$ are constant matrices and p_l are the scheduling variables. The dependence can also be modeled as not affine, e.g., linear splines [8, 13] and polynomials [10, 12]. The most common way to model the dependence on the scheduling variables is static, i.e. the parameters in the system matrices are only dependent on the current value of p. If there exists dependence on past values of p, the dependence is called dynamic [5, 6].

Previously used LPV models for the Li-ion battery are divided mainly in two ways: what scheduling variables are used and how the dependence on the parameters are structured. The previously used scheduling variables are temperature, SOC, SOH, and current direction. The most common scheduling variable is temperature, which is reasonable considering its widely known effect on the resistance of the battery [21]. Temperature is also easily measured, in contrast to SOC and SOH. The scheduling variable dependence vary between simple linear dependence to more complex exponential or polynomial functions.

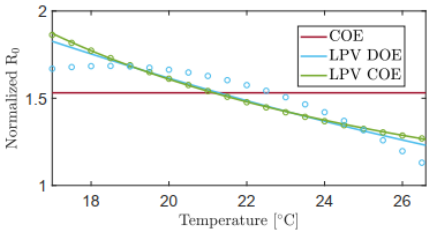
As stated earlier an affine dependence on the scheduling variables is a common way to structure the LPV model. This is used in [9], where both temperature and SOH are used as scheduling variables. This gave promising results, but what has to be considered is that this study focused on hybrid EVs and could thus neglect the effect of SOC due to the limited operating range. The method was also developed for an on-board application for diagnosis of SOH, which sets it apart from this project.

The studies from [8, 13] both use SOC as the scheduling variable, but do discuss the expanding of the model to include temperature. The interpolation between models is done with linear spline functions, that are easily extended to more than one dimension. The study in [8] found that this type of model provided good results in a limited operating region of SOC, excluding high and low values of SOC.

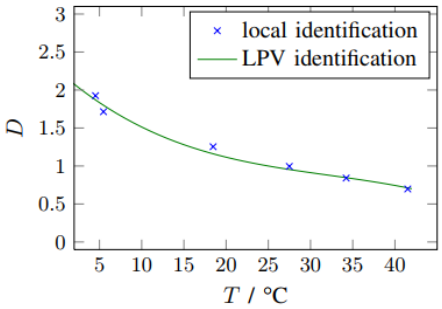
In [10, 12] similar models are used where temperature is the scheduling variable. The input u to the model is extended from a scalar to a vector to include elements that take the form IT^n where I is the current, T is the temperature and $n \in \mathbb{Z}$. The scheduling functions thus takes a sort of polynomial form, since powers of the scheduling variable is introduced to the input. For [12], which uses an output-error model, this requires that the model is extended with more transfer functions that model the dependence on each input element in the extended input vector. By structuring the model like this, the parameters of the model need no temperature dependence. Despite this polynomial structure, [10] did find that the dependence on temperature for the A matrix was close to linear. The B and D matrices did not however have this simple linear behavior, indicating the need for the more complex polynomial based model structure. The dependence on temperature for the D matrix, i.e. the model parameter R_0 are similar for both [12] and [10] and can be seen in Figure 2.5.1. The figures show the expected decrease in resistance as the temperature increase. The dependence is captured with a quadratic dependence on the form $R_0 = \beta_0 + \beta_1 \frac{1}{T} + \beta_2 \frac{1}{T^2}$ in [12] and with a spline function in [10], illustrating how the dependence can be described in different ways.

Even though it is different from the other LPV structures discussed, the study in [11] offers an interesting approach that only uses current direction as the scheduling variable. The parameters in this model are thus binary quantities, one for charging current and another for discharging. This study however focuses on an isothermal setting that consequently disregards the well known effect of temperature on the battery parameters.

In [14] a model is implemented that uses a function of the state and input as the scheduling variable, thus making it a quasi-LPV model.



(a) The temperature dependence of the model parameter R_0 matrix found in [12] plotted as green points with a well fitted green quadratic function labeled (LPV COE). The red line (COE) represents a constant model and the blue points and fitted function (LPV DOE) are from a discrete time model that did not give satisfactory results.



(b) The temperature dependence of the D matrix found in [10] with the fitted spline function also plotted.

Figure 2.5.1: The ohmic resistance R_0 as a function of temperature from two different studies.

2.6 Identification techniques

Several different identification techniques have been used to identify the LPV models. In [9, 12] least squares algorithms are used to find the parameters, which is a common way of structuring system identification. Another common identification method is subspace identification, which is used in [10]. Subspace identification has the added benefit of providing a suggested model order, which for the case in [10] was a model of order 2.

In [8, 13] a genetic algorithm is used to identify the LPV model. The idea with the genetic algorithm is that it simulates a natural selection process. The identification procedure starts with a set of randomly parameterized models which are all evaluated on the dataset. The best models are then paired up and combined to a new model, with some slight random mutation. The models that performed worst in the evaluation are discarded. Under the right conditions this algorithm then converges to a minimum after a number of iterations. Using a genetic algorithm might be attractive because of the highly automated procedure of finding all the parameters in one unified process. One must however take care to ensure that the algorithm will work. Physically reasonable bounds must be set on the parameters to ensure that the solution converges to a minimum. The

identification scope must also be gradually increased, starting with the identification of only some smaller set of the complete set of parameters, and then including the rest of the parameters in the identification process. Even with these optimizations of the process, [13] reported a 4 hour identification procedure.

The System Identification Toolbox in MATLAB [22] is a set of tools that aids system identification of a variety of different kinds of continuous- and discrete time models from given input-output datasets. It also offers tools for evaluating and comparing models against each other. By setting constraints on the identification procedure the toolbox can also be used for gray-box system identification. The models identified with the toolbox are also easily integrated with Simulink. Another toolbox that can be used for parameter optimization is the Simulink Design Optimization Toolbox [23]. It provides tools for analyzing and tuning model parameters within the Simulink environment. It also lets the user investigate the model's sensitivity to changes of each parameter, which is highly relevant for this project since accurate model parameters are key to a good model. MATLAB and Simulink tools were used in studies for battery parameter estimation in [11, 24–26] with promising results. The System Identification and Simulink Design Optimization toolboxes offer simple, but powerful and fast tools for system identification and are thus suitable tools for this project that will handle many separate identifications.

2.7 Drive cycles

The RPTs offer a good hint of the battery behavior, but are carried out under very controlled circumstances and might thus not capture battery behavior during real use. In order to optimize the model using data more reflective of real battery behavior during driving one can use drive cycles (DC). DCs offer measurements from the battery just as the RPTs, but the charging and discharging is caused by real use of the battery in an EV, rather than the perfectly shaped pulses seen in RPTs. DCs will also naturally have a variation in the possible scheduling variables as SOC, temperature and current changes during driving.

Chapter 3

Method

In this chapter the general model structure is described. The identification procedure consisting of first the local, and then the global approach is provided. Lastly the methods of validating the LPV model is specified.

The method of the project can be divided in to two parts. The first consists of identifying local models at fixed values for the possible scheduling variables. These identifications lay the foundation for finding possible scheduling functions. From this choice of scheduling functions a global estimation can be set up to optimize the parameters of the scheduling functions. The result of the global estimation will then be the final LPV model. As this project has a limited time span the choice was made to limit the project to a first order ECM. This will limit the accuracy of the model, but will lower the complexity of the identification procedure and reduce the number of parameters of the final model. This means that the final system is described by the state space description

$$\begin{aligned} \dot{x} &= -\frac{1}{R_1(p)C_1(p)}x + \frac{1}{C_1(p)}u \\ y &= x + R_0(p)u, \end{aligned} \tag{3.0.1}$$

where the state x is the voltage over the RC network, $R_1(p)$, $C_1(p)$ and $R_0(p)$ are the scheduling functions dependent on the scheduling variables p. The tools used for the identifications are part of the MATLAB and Simulink environments. The studies in Section 2.4 suggest that temperature, SOC, current and SOH are possible scheduling variables. What limits this project is the lack of available data for different levels of SOH. This means that the effect of SOH will be hard to both identify and validate. Thus, this project will focus on temperature, SOC and current as possible scheduling variables.

3.1 Local identification

The local identification consists of first estimating the circuit parameters for each operating condition, i.e. all possible values of the scheduling variables. From this, the possible scheduling functions can be designed.

3.1.1 Circuit parameter estimation

The local identifications are performed using HPPC RPTs, as described in Section 2.3. These tests are picked out so that the variation in the possible scheduling variables is as big as possible. The operating conditions in these different RPTs will then span the scheduling space. Since the HPPC tests are measurements taken directly from the battery cell, the OCV must first be removed in order to acquire the dynamic part of the cell voltage as described in Section 2.1.1. By the use of already established SOC-OCV curves the OCV can be calculated for the duration of the test and then removed from the terminal voltage. In order to ensure that the identification converges to the right minimum, the method for acquiring an initial guess for the circuit parameters described in [18] is used. Having this initial guess of the circuit parameters, one can then construct the initial guess for the system matrices in 3.0.1. These initial guesses are used when the estimation is done using the System Identification Toolbox in MATLAB. With the circuit parameters estimated, the goal is then to relate them to the possible scheduling variables.

3.1.2 Scheduling functions design

When all of these separate system identifications has been done, the parameter dependence can be studied in order to find suitable scheduling functions. By plotting the circuit parameters for different values of the scheduling variables, trends can be seen. As a secondary step, the Curve Fitting Toolbox in MATLAB [27] provides a convenient way of importing data to be used for fitting different kinds of functions. This makes it a suitable tool for finding possible scheduling functions. The fitted functions can then be exported to the MATLAB workspace to be used in other scripts or Simulink files. Polynomials offer a good foundation for functions to test since they are easy to increase and decrease in flexibility by adding or removing terms. A key to finding a model with as few parameters as possible is also to investigate the importance of each polynomial term and remove terms with smaller magnitude, since these will not affect the function value as much as terms of greater magnitude.

The Simulink environment provides all the necessary tools for setting up a LPV model and simulating its output given some input and measured scheduling variables. The simulated response can then be examined directly in Simulink or exported to the MATLAB workspace. By importing the candidate scheduling functions and simulating the response using these functions, the model performance can be evaluated using the fit described in 2.1.3, and a set of suitable models for further investigation can be found.

3.2 Global identification

The scheduling functions from the local identification will form a complete LPV model. The model parameters can however be optimized by using a dataset that spans the scheduling space well. For this project drive cycles (DC) provide fitting data for this. In order to ensure that the optimized model is not biased to a certain operating region, the DCs must have a range of the scheduling variables that is as broad as possible.

The Design Optimization Toolbox in Simulink [23] has a tool for parameter estimation. In the tool the experiment with input and output data can be chosen along with the parameters that should be optimized. The parameters are adjusted so that the error between the measured and simulated output is minimized until some convergence criteria is met or a maximum number of iterations has passed. Since the Parameter Estimator App can take a long time until it reaches the convergence criteria, it will be necessary to find a smaller number of interesting models to investigate with the global identification. Investigating too many models will thus take too long, and the time limitation of the thesis will not hold.

3.3 Evaluation

When a number of candidate models result from the global estimation, some sort of evaluation needs to be carried out to asses the performance of the models. This evaluation will be comprised of two aspects, accuracy and complexity. The accuracy is fairly straight forward and can be validated through simulations where the candidate models are fed with the same input, and the outputs are compared to the real output though the fit measure from 2.1.3. This evaluation should also be done on a different data set than the one used for the parameter optimization so that overfitted models are not favored. The aspect of complexity is most easily described by the number of parameters in each model. It is expected that models with a higher number of parameters in general will perform better, but this better performance must be weighed against the larger number of parameters and thus increased complexity. Another part of the evaluation process will be to compare the LPV model to the currently used model in the BMS at Scania, hereby called the Scania model. What is relevant for this comparison is naturally the accuracy of the model in terms of model output compared to measured output. What also is highly relevant is the implementation cost of the LPV model compared to the Scania model in terms of required computer power. This comparison has many layers to it, but one valid measurement would be the number of calculations that each model requires to simulate a given output sequence.

Chapter 4

Results

In this chapter the details of the construction of the LPV model based on the found parameter dependencies is described. The choice and fitting of the scheduling functions is specified. From this, the global optimization using the DCs is described, along with the validation of both the unoptimized and optimized LPV models.

4.1 Local identifications

The first step of the local identifications was to get initial guesses of the circuit parameters with the method described in [18]. This was done by programming an algorithm that picked out a single pulse and following relaxation period in the HPPC RPT and extracted the circuit parameters from that. These values were then used as the initial guess for the model estimation done with the **ssest** function in MATLAB. While the initial guess for the circuit parameters was done with just one of the pulses, the model estimation used the full length of the HPPC RPT.

4.2 Parameter dependence

The first dependence that was investigated was the temperature dependence. The available data sets had HPPC tests in a range between -18° C to 38° C with increments of about 10° C. The results of the temperature dependence for the parameters can be seen in Figure 4.2.1. In the figure it can be seen that the resistance parameters R_0 and R_1 shows the expected increase in resistance as the temperature decreases. The shape of these curves is also visibly non-linear. The capacitance C_1 does however show an apparent linear increase in capacitance as the temperature increases.

4.2.1 Temperature and SOC dependence

As described in Section 2.4 it is also known that SOC greatly affects the circuit parameters. The circuit parameter dependence on SOC in combination with temperature was thus also investigated. The available HPPC test had SOC in the range between 10% to 80%. The parameters dependence on both temperature and SOC can be seen

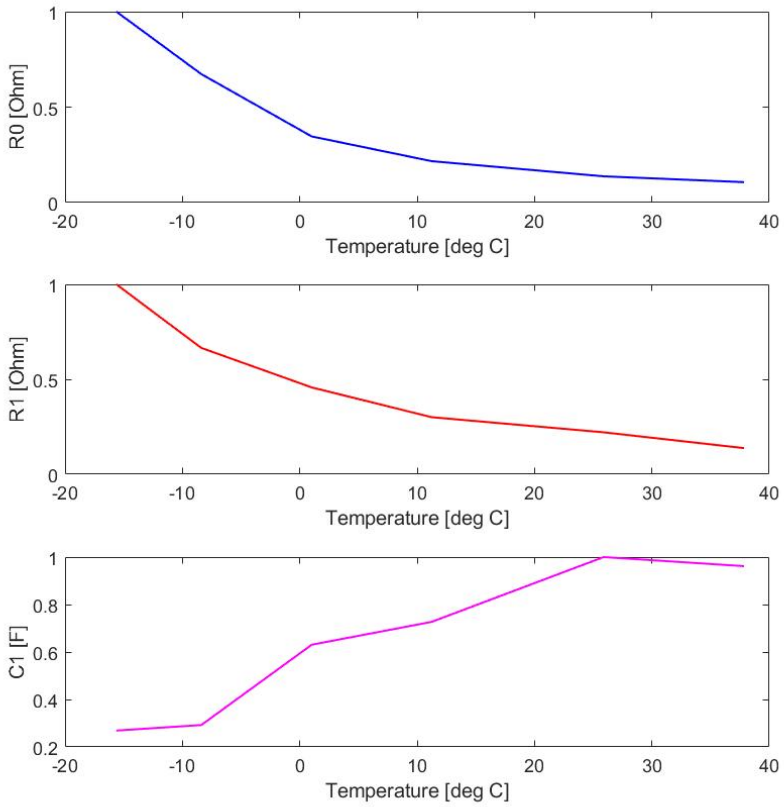
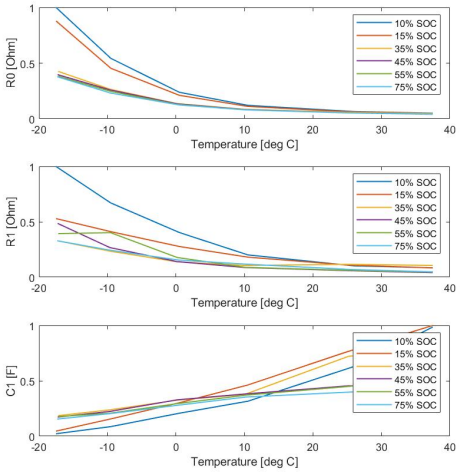
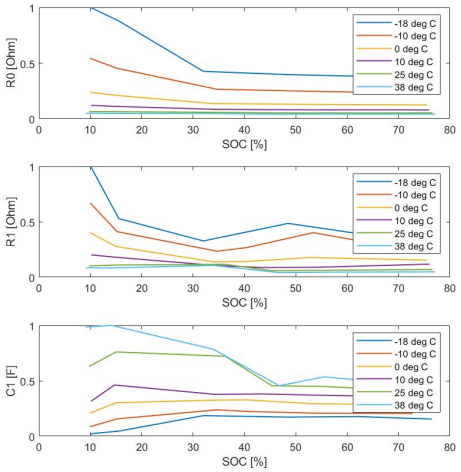


Figure 4.2.1: The circuit parameters of a first order ECM and their dependence on temperature. The axes for the circuit parameters have been normalized.



(a) The model parameters dependence on temperature for different SOC.



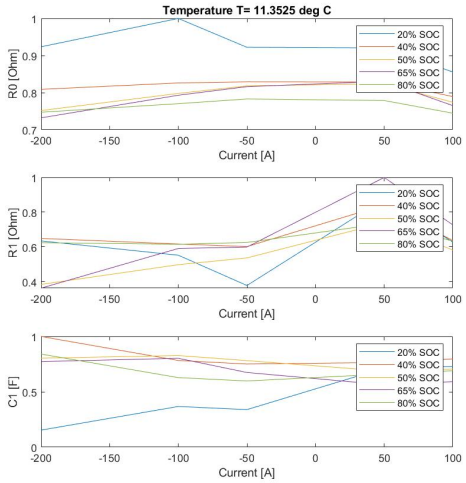
(b) The model parameters dependence on SOC for different temperatures.

Figure 4.2.2: The model parameters dependence on both temperature and SOC. The axes for the circuit parameters have been normalized.

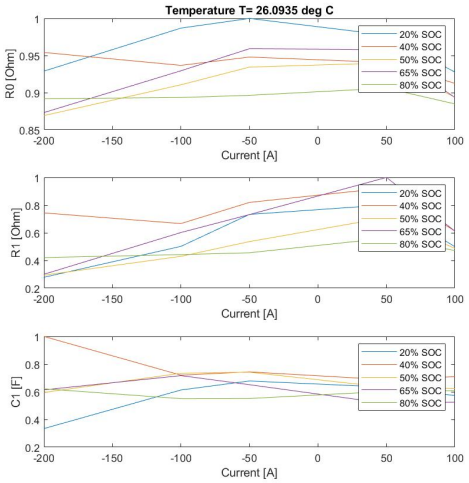
in Figure 4.2.2. In Figure 4.2.2a a similar trend for the resistance parameters R_0 and R_1 as in Figure 4.2.1 can be seen. The slope of the curves is however increased as SOC decreases. This increase of the slope is most prominent for SOC below 30%, as seen in Figure 4.2.2b. The capacitance C_1 shows the apparent linear behavior described earlier. The capacitance does not show any apparent functional behavior with respect to SOC, as seen in 4.2.2b. One possible behavior could be that the slope of the curve is increased as SOC is decreased.

4.2.2 Current dependence

The dependence on current is only used in previous studies in term of dependence on current direction. The study in [21] does however suggest that the resistance parameters have a clear dependence on current magnitude. This dependence was thus also investigated in order to find any meaningful dependence. In Figure 4.2.3 the dependence on currents ranging between -200A and 100A can be seen. The figures include plots for both different SOC and temperature. In Figure 4.2.3a, where the temperature is around 10°, the trend that higher current magnitude results in a lower resistance can be seen for the R_0 parameter. This dependence is not as strong as in Figure 2.4.2. In Figure 4.2.3b, where the temperature is around 26°, the dependence on current seems negligible. In Figure 4.2.3b it can be seen that for charging current the resistance R_1 is higher, and that this effect is increased as SOC decreases. This trend is not however apparent in Figure 4.2.3a. No clear dependence can be seen for the capacitance C_1.



(a) The model parameters dependence on current for different levels of SOC. The temperature here was about 10°.



(b) The model parameters dependence on current for different levels of SOC. The temperature here was about 26°.

Figure 4.2.3: The model parameters dependence on current for different levels of SOC and temperature. The axes for the circuit parameters have been normalized.

4.2.3 Choice of scheduling variables

Conclusively it can be said that temperature and SOC affect the model parameters more than current. Due to this, the scheduling variables used for the LPV model will be temperature and SOC. These factors show a clear effect on the model parameters. This result is also in line with the previously studied LPV models for the Li-ion battery, where temperature and SOC were the most commonly used scheduling variables. As the trend in Figure 4.2.2 shows how the circuit parameters seems to converge to some more or less fixed value for high SOC, the choice was made to copy the circuit parameter values for the highest level of SOC measured to 100% SOC.

4.3 Scheduling functions

The polynomial based scheduling functions were constructed using the Curve Fitting Toolbox in MATLAB. As two scheduling variables were to be tested, the polynomials included cross terms with both temperature and SOC. These polynomials were thus surface polynomials of the form

$$f(z,T) = p_{00} + p_{10}z + p_{01}T + p_{11}zT + p_{20}z^2 \ldots, \tag{4.3.1}$$

where $f(z,T)$ is the scheduling function for R_0, R_1 or C_1, z is the SOC, T is the temperature and p_{xy} is the coefficient for the term of degree x in SOC and degree y in temperature. In Figures 4.3.1 and 4.3.2 two examples of these polynomials can be seen. A reasonable upper limit for the degree of the polynomial in each dimension was 3. This upper limit was set partly to restrict the number of parameters, which is important for both model

Degree in SOC / Degree in temperature	1	2	3
1	67.981	69.767	70.032
2	62.522	70.193	73.504
3	58.169	72.797	73.240

Table 4.3.1: fit results for scheduling functions of equal form.

complexity and the limited scale of the thesis. It is also important to keep the number of parameters down to simplify the global identification procedure. To get an initial overview of what polynomial structures were interesting for optimization in the global identification, all combinations between degree 1 and degree 3 were fitted to the locally identified circuit parameters. Different combinations of these scheduling functions were then used in the Simulink LPV model. A DC was used as validation data and for each tested combination the fit of the output was noted. From this, a first general selection of interesting polynomial structures could be made. The first approach was to let all the scheduling functions be of the same order. These choices of scheduling functions were then tested on the DC. The fit results from these tests is found in Table 4.3.1. What can be said in general is that higher order scheduling functions give better results. More specifically, increasing the degree in temperature from 1 to 2 gives a big improvement in performance. What also can be seen is that for degree 1 in temperature, increasing degrees in SOC decreases the model performance. The best choice of scheduling functions according to this test was functions of order 2 in SOC and order 3 in temperature. A modified version of this model was also investigated further. The modification was to remove the parameters of least importance in the R_0 and R_1 scheduling functions. This was done by examining the magnitude of the coefficients for the terms in the polynomials, with consideration taken to the effect of the exponentiation, and removing the terms with smallest coefficient magnitude. As seen in Figure 4.2.2, the dependence on SOC for C_1 is limited. The same kind of complete test of scheduling functions was thus made with a fixed scheduling function for C_1 with degree 0 in SOC and degree 1 in temperature. These results are found in Table 4.3.2. The fit values are very similar to the ones in Table 4.3.1, indicating a mild effect in general of the C_1 function. As an additional investigation the same test was done without SOC dependence in all scheduling functions. This gave similar results for all degrees in temperature. Both degree 1 and 2 in temperature were thus investigated further. As the general results of testing different kinds of polynomial orders of scheduling functions did not point to one especially good choice, some additional configurations of scheduling functions were also investigated further to get a better overview. As a reference, a constant scheduling function model was also included in the set of candidate models. Since this model does not have a dependence on the scheduling variables it is not LPV, and just a LTI model. All the models that were tested in the global optimization is listed in Table 4.3.3. Model D has the same degrees of the scheduling functions as model C, but this model has the earlier described removal of unimportant terms in the polynomials.

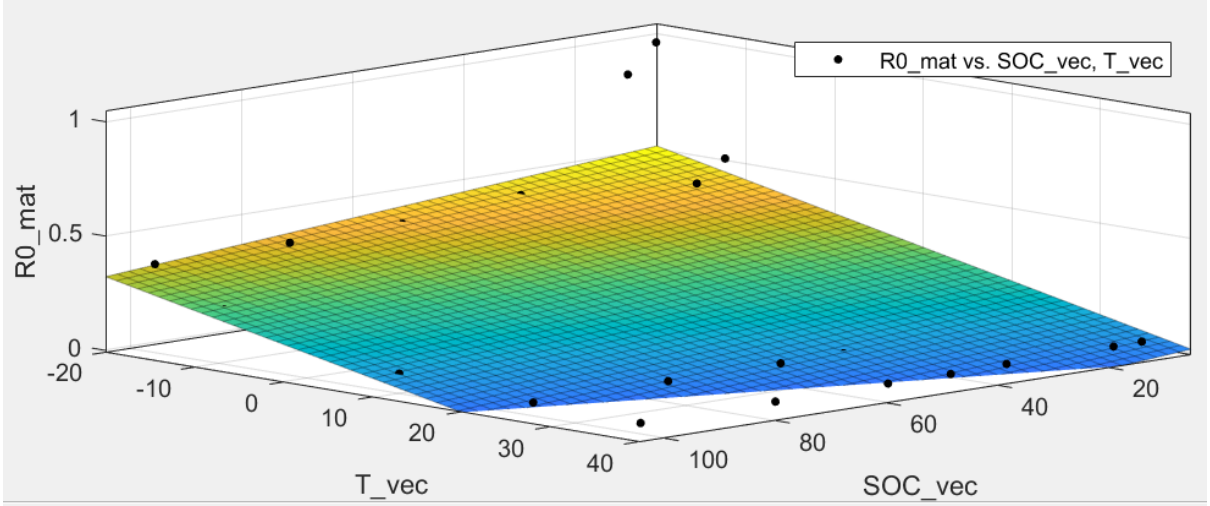


Figure 4.3.1: A scheduling function for the circuit parameter R_0 with degree 1 in both SOC and temperature. *T_vec*, *SOC_vec* and *R0_mat* refers to temperature SOC and R_0. The R_0-axis has been normalized.

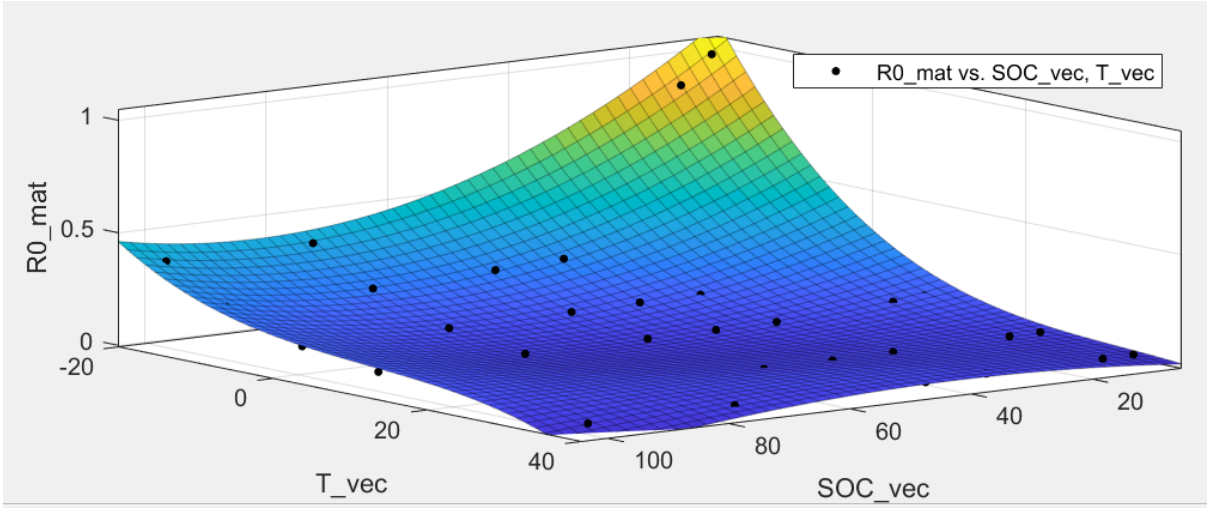


Figure 4.3.2: A scheduling function for the circuit parameter R_0 with degree 2 in SOC and degree 3 in temperature. *T_vec*, *SOC_vec* and R0_mat refers to temperature SOC and R_0. The R_0-axis has been normalized.

Degree in SOC / Degree in temperature	1	2	3
1	68.073	69.900	69.858
2	62.278	70.278	73.223
3	58.713	73.003	73.210

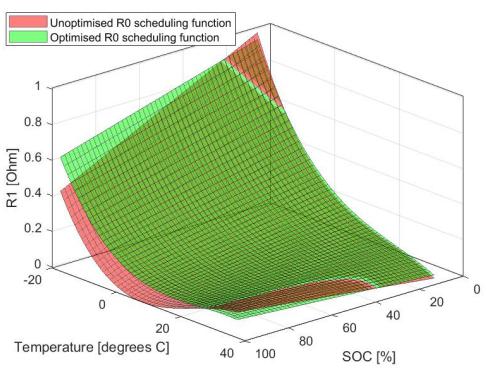
Table 4.3.2: fit results for different scheduling functions for R_0 and R_1, but constant scheduling function with degree 0 in SOC and degree 1 in temperature for C_1.

Model	Degree in SOC for R_0 and R_1	Degree in temperature for R_0 and R_1	Degree in SOC for C_1	Degree in temperature for C_1	Number of parameters
A	1	1	1	1	9
B	0	2	0	1	8
C	2	3	0	1	20
D*	2	3	0	1	14
E	1	1	0	0	7
F	0	1	0	1	6
G	1	2	1	1	13
Constant	0	0	0	0	3

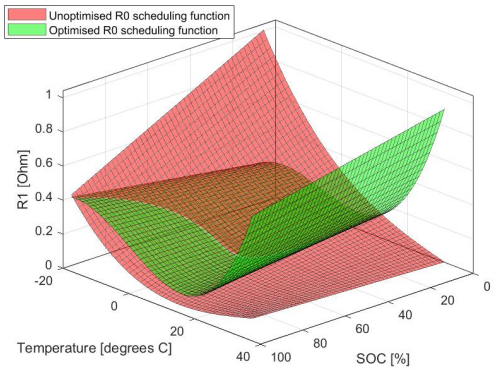
Table 4.3.3: The model candidates for the global optimization.
* This model is the modified version of model C with the least important terms removed.

4.4 Parameter optimization

As a first approach the parameters of the model candidates were optimized in the Design Optimization Toolbox using one single DC. This DC started at a temperature of -20° C and about 40 % SOC, and ended on 0° C and about 15 % SOC. When the optimized models were validated on other DCs the validation fit was low, and in some cases negative. The resulting models were evidently too biased to the DC used for optimization. This made it apparent that more than one DC was needed for the global optimization. The new approach used four different DCs with varied scheduling trajectories. Two additional DCs were reserved for evaluation. The scheduling trajectories of these six DCs can be seen in Figure 4.4.1. In the figure it can bee seen that the four DCs used for optimization span different regions of the scheduling space. The DCs were loaded in to the Parameter Estimator App and the optimization algorithm worked to minimize the summed squared error of the four DCs. The algorithm ran until the convergence criteria was met. This criteria was dictated by the parameter tolerance that was set to 0.0001. The Design Optimization Toolbox has a variety of different methods and algorithms to pick between. The one used for this project was the default setting with nonlinear least squares as method and trust-region-reflective as algorithm. When the optimization converged, the new parameter values were saved. This was done for each of the models A-G in Table 4.3.3 and the optimized parameters were saved. An example of the R_0 and R_1 scheduling functions before and after optimization can be seen in Figure 4.4.2. In the figure it can be seen that the unoptimized scheduling function in Figure 4.4.2a dips very close to zero for high SOC and temperatures around 10° C. The optimized scheduling function does however not do this, indicating a possible benefit of the optimization procedure. The scheduling function R_1 in Figure 4.4.2b changes quite drastically after the optimization. The unoptimized and optimized version of the scheduling function are close to each other in the regions of the scheduling trajectories for the DCs used for the optimization, but for other values of SOC and temperature, the optimized scheduling function is diverging from the unoptimized.



(a) The scheduling function R_0 for model D before and after the optimization of the model parameters.



(b) The scheduling function R_1 for model D before and after the optimization of the model parameters.

Figure 4.4.2: The scheduling functions R_0 and R_1 of model D before and after optimization. The functions have been normalized.

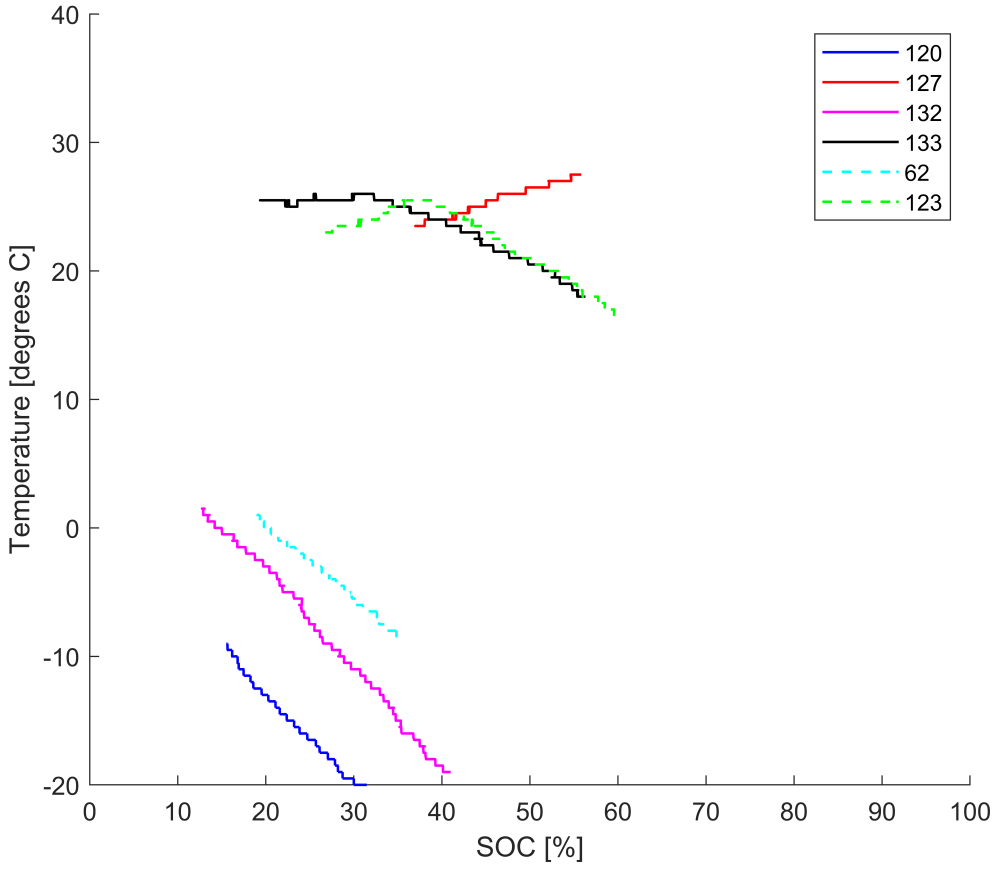


Figure 4.4.1: The scheduling trajectories of the DCs used for optimization seen in full lines, and the ones used for validation seen in dashed lines.

Model	Validation on DC 123 before optimization	Validation on DC 62 before optimization	Validation on DC 123 after optimization	Validation on DC 62 after optimization
A	74.3789	81.8945	66.6153	82.0407
B	30.7195	74.4165	78.1056	83.0943
C	61.1974	74.7198	81.1508	83.6431
D	70.4958	82.1904	77.8447	85.1151
E	74.5941	81.4071	64.9319	78.6853
F	67.3965	84.4366	61.0771	83.0178
G	30.7836	85.7570	78.7107	84.1089
Const.	52.5454	34.8897	75.4194	26.5692

Table 4.5.1: The model candidates validation fit on the two different driving cycles before and after the parameter optimization.

Model	DC 123	DC 62
Scania model	76.4583	75.5165
Best LPV model	81.1508	85.1151

Table 4.5.2: The Scania model validation fit on the two different driving cycles and the validation fit for the best performing LPV model for each validation DC.

4.5 Model validation

The optimized models A-G were all tested on the driving cycles 123 and 62. The scheduling trajectories of these DCs are quite different from each other, as seen in Figure 4.4.1. The fit results of the validation is listed in Table 4.5.1. The results show that for DC 123 the fit was better for all models (constant model not counted). For DC 62 the results are better than for DC 123, but the trend is generally the opposite, with better fit for the unoptimized model for all but two models. This might be explained by the fact that the scheduling trajectory for DC 62 lies somewhat apart from that of 120 and 132, while 123 is close to 127 and 133. What also can be noted is that the unoptimized models validation fit on DC 123 is more uneven than that of DC 62. The validation on DC 123 does show that the models that had an order of 2 or higher for the degree in temperature in scheduling functions R_0 and R_1 is notably better than the models with a linear dependence on temperature. The overall best model was model C after optimization, having validation fits over 80% for both validation DCs.

In Table 4.5.2 the fit on the validation DCs can be seen for the Scania model. The validation fit of the Scania model is better than some LPV models and worse than some. The best LPV models are however considerably better than the Scania model for validation on DC 62.

4.6 Cost of implementation

The LPV model has at most 20 parameters, which means that this gives a static memory requirement of at most 20 constants. What needs to be added to this is the static memory requirement caused by the storing of the functions that returns the parameters. For the LPV model these functions ranges between linear functions, to polynomial functions of degree three, with varying degrees of polynomials in between. The LPV models' dynamic memory requirement will be based on the complexity of the scheduling functions. In Figure 4.6.1 the functions' calls required for one simulated time step are plotted against the number of model parameters in each model A-G. In the figure it can be noted that there is a close to linear relationship between the number of model parameters and the number of function calls.

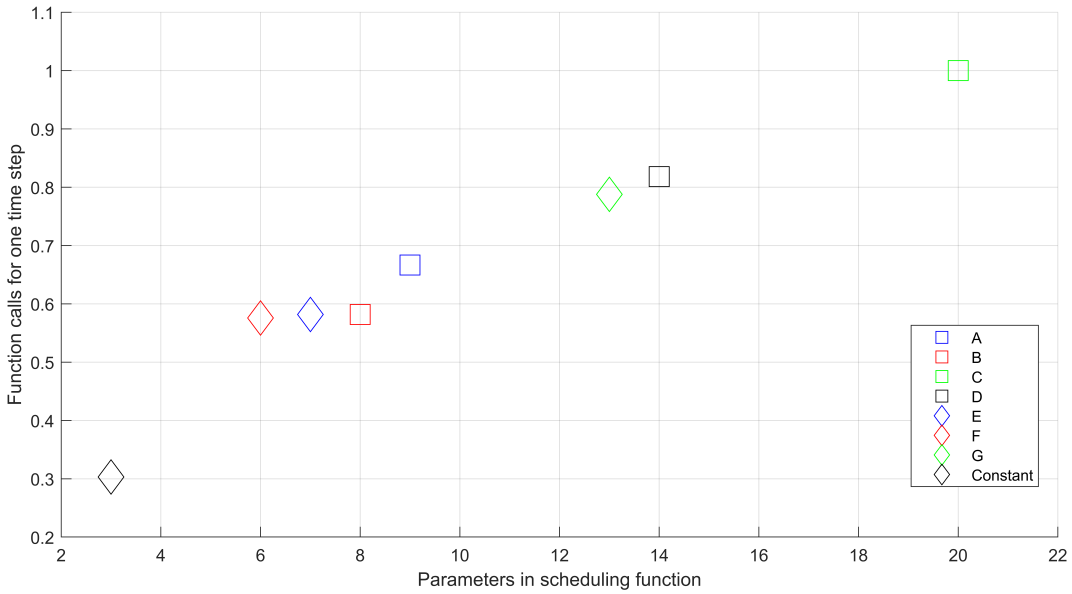


Figure 4.6.1: The number of parameters in each LPV model and the number of function calls made for one simulated time step. The function call values have been normalized.

Chapter 5

Discussion

In this chapter the results are discussed in terms of model performance and validity. The discussion is focused on the limitation of the DCs, issues with the SOC and OCV, and lastly alternative LPV structures and purposes are described.

5.1 Effects of DCs

It is apparent from the results that the global approach results did not entirely meet the expectations that the validation performance would be improved after the optimization. This would suggest that the benefit of doing the global optimization is not very clear. Add to this the extra work of setting up the global optimization and the case for opting for the local approach becomes even stronger. What however must be taken in to consideration is the location of the validation DCs 123 and 62 relative to the DCs 120, 127, 132 and 133 used for identification. As discussed in the results, the validation with DC 123 that achieved generally better results after the global optimization lies very close to DC 133 in the scheduling space, as seen in Figure 4.4.1. For the validation with DC 62 that had generally better results before the global optimization it can be noted that its scheduling trajectory does not coincide with a DC used for optimization. Another complicating factor is that the DCs are of different lengths in time. This means that if DCs of different length are used for optimization there will be a bias towards the longer DC since the Parameter Estimator App aims to reduce the summed squared error. This issue could possibly be resolved by resampling the DCs so that they all contain the same number of samples, regardless of length in time.

Looking at Figure 4.4.1 it is apparent that the general trend of the DCs is that they go from a low temperature and high SOC, to a high temperature and low SOC. This shape of the DC is expected since driving will both lower the SOC, and heat up the battery, increasing the temperature. More specifically, any change in SOC is expected to increase the temperature since change in SOC means current is flowing either from or to the battery, heating it up. It is obvious that the coverage of the scheduling space needs to be better. The currently used DCs only covers two general regions. Ideally, the whole scheduling space should be covered by having a greater spread of the DCs and more DCs than what is currently used. The drawback with having more DCs is that the optimization procedure becomes more complex and time consuming. One way to

find a compromise between coverage and complexity could be to find, or acquire through drive tests, DCs that cover strategic parts of the scheduling space. By doing this the optimization could still be quick and overfitting is avoided.

The discussed issue of the optimized models being biased towards certain regions of the scheduling space might be an explaining factor in how low order linear models performed relatively good. If the optimized scheduling functions only needs to capture the behavior of two general regions it is reasonable that a lower order model could achieve this. What also contributes to the relatively good performance of the low order models might be the position of the DCs used for optimization. In Figure 4.3.2 it can be seen that the most extreme region of the scheduling space is at low SOC and low temperature. This region is covered neither by the DCs used for optimization nor by the DCs used for validation. The unintentional exclusion of this region of the scheduling space is likely another explaining factor in why the lower order models perform so well. What also must be addressed in the results from Table 4.5.1 is that model F, that does not have any dependence on SOC, shows a reasonably good validation fit. This would suggest that SOC is not a necessary scheduling variable. This is yet another effect of the limited coverage of the DCs since the whole scheduling space should be covered to see the benefit of using SOC as a scheduling variable. This is also backed up by the results from earlier studies in Section 2.4 that concluded that, for a new battery, the effect of SOC is limited to only very high or low SOC. Another drawback of limited coverage of the scheduling space is the overfitting that might occur for the more flexible scheduling functions. This is what can be seen in Figure 4.4.2b. This is naturally unwanted behavior that must be corrected if the LPV-structure were to be used in the BMS.

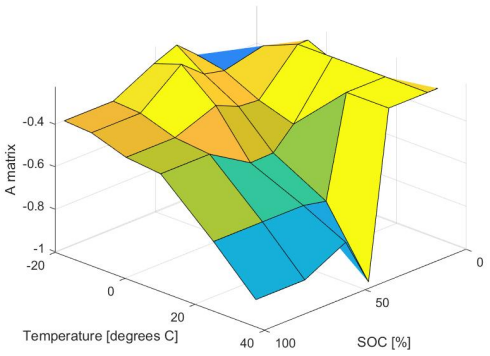
5.2 Issues with SOC and OCV

As discussed in Chapter 2, a possible purpose of the battery model is to use it for SOC estimation by removing the polarization voltage simulated from the model from the terminal voltage to acquire the OCV. A potential problem with this might be that the SOC is used in the model as a scheduling variable. This could in some situations create an undesirable feedback loop that could affect the model output drastically.

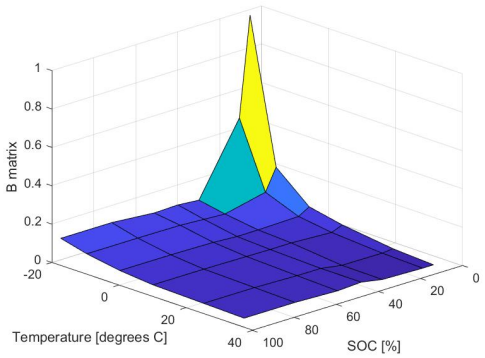
A crucial part of both the local identification and the global optimization is the removal of the OCV from the measurements of the terminal voltage. This proved to be somewhat complicated in some situations where the measurements did not have a stable and relaxed starting voltage, since that is needed to estimate the starting SOC. The OCV is several orders of magnitude larger than the polarization voltage so, a small error in the OCV will therefore shift the model output relatively far of from the measured output and greatly lower the fit. This error in the calculated OCV can either be caused by an incorrect staring voltage, or a mismatch in the capacity used for the calculation of the SOC. For a RPT this mismatch is not likely to cause a big error in the SOC, but for a several hours long DC, it can have an effect.

5.3 Alternative LPV structures

The benefit of having the scheduling functions be connected to the circuit parameters is that the physical connection to the gray-box model is kept. This kind of model lets one see how the resistance and capacitance changes with SOC and temperature and relate this to previously known dependencies. The drawback with this model structure is that the model is not linear in the parameters. The way the scheduling functions are multiplied with each other and then put as the denominator in a rational function, that is not possible. The point of having the model be linear in the parameters is that the global identification procedure could be done with a simple least-squares method, rather than the currently used non-linear method in the Parameter Estimator App of the Design Optimization Toolbox. The Parameter Estimator App is simple to use and works well, but the optimizations for the higher order models take up to an hour to converge. A direct least-squares identification of the model would be much faster as it does not involve an iterative procedure. A simple approach to set up a model like this would be to have the system matrix elements be of the same type of polynomials that is used for the current scheduling functions. This would likely require different degrees of the polynomials because of the different dependencies that the system matrices have on the scheduling variables. The A- and B-matrices dependency on SOC and temperature is illustrated in Figure 5.3.1. In the figures it can be seen that the shapes of the surface plots likely would require more flexible functions to be described well.



(a) The system A-matrix dependency on the scheduling variables.



(b) The system B-matrix dependency on the scheduling variables..

Figure 5.3.1: The system matrices dependency on the scheduling variables. The matrix values are normalized.

A possible implementation of the LPV-model structure studied in this thesis could be to have it be part of an online model. While an offline model, like the one studied in this thesis, has fixed parameters estimated before use of the model, an online model continuously estimates the parameters based on the measured input and output data to fit the current operating conditions. A key benefit with an online model structure like this would be that it could adapt to the changing battery behavior as the battery ages. Since the current LPV model is identified offline on data from beginning of life, it will perform worse as the battery ages. An online model structure could possibly also use a

lower order linear model without having the risk of not capturing extreme behavior, such as what is seen for the resistance at low SOC and low temperature. As the scheduling variables typically change slowly, the online model could have a slow update rate on the parameters. A slower update rate is beneficial as it reduces the required computing power. If an online model structure is to be implemented, it will however be necessary to construct the scheduling functions differently so that the system matrices are linear in the parameters, so the identification can be computationally cheaper.

Chapter 6

Conclusions

In this chapter the conclusions of the thesis is provided based on the results and the discussion. Possible further work on similar projects to this is also described.

The fact that there are benefits with using an LPV model compared to a constant LTI model is clear. The LPV models are able to capture the changing battery behavior as the temperature and SOC change. Whether the global optimization is necessary is not obvious. As the validation results show both better and worse models after optimization it is hard to make a definite strong case for the global optimization. Although it can be concluded that the best overall model was model C *after* optimization. As discussed, there are however several aspects of the global optimization that likely would improve the procedure. These improvements should also give results that are even more in favor of the higher order and more flexible models. An additional indication that the higher order models have a greater potential is that these models in general performed better after the optimization procedure than before, while the lower order models, with a fewer number of total parameters, performed in general worse after the global optimization. Another benefit with the global approach is that it is quite effortless. When the structure of the LPV model has been decided the optimization procedure is fairly straight forward with the use of the Design Optimization Toolbox, even if it might require some time. To set up the local model one needs to conduct or find tests taken at the various combinations of SOC and temperature, and structure the system identification on all of these tests.

The validation results show fit values in the range of about 60-85% which is not excellent seeing how LPV models from previous works achieved higher values. What has to be taken in to consideration is that this model is a first order model, while the previous works in general use higher order models. The best LPV models did however outperform the currently used model at Scania, which makes a strong case for the LPV model.

6.1 Further work

As this thesis is limited in many aspects, largely due to the limited time of one semester, there are many parts of the project that can be studied further. The first and most urgent improvement would be to sort out the issue with limited DC coverage of the scheduling space. This is largely a matter of taking the time to sort through available data and put together relevant combinations of DCs to use for the global optimization.

This project focuses on polynomials as scheduling functions, but this choice is largely driven by the need for a simple function structure that is easy to scale up and down in flexibility, and thus, number of parameters. Further works similar to this thesis should explore more complex function structures. As described in Section 2.4 it is known that the resistance has an exponential relation with temperature. This is an example of knowledge that could be used. A key aspect of using other types of functions could also be to have functions that naturally do not attain negative values, such as an exponential function. Though it is expected that R_0 and R_1 have similar dependence on the scheduling variables, it might also be an option to use different scheduling function structures for these circuit parameters to reduce the number of parameters in the model.

The above described improvements are fairly straight forward and do not require large restructuring of the work process of this thesis. Bigger changes than these things could be to use a second order model, include the effect of hysteresis or investigate the effect of a dynamic scheduling dependence.

Bibliography

[1] IEA, *Global EV Outlook 2020*, https://www.iea.org/reports/global-ev-outlook-2020. Accessed February 19, 2020.

[2] Scania, *Electrifcation of an industry*, https://www.scania.com/group/en/home/about-scania/innovation/technology/electrification.html. Accessed February 19, 2020.

[3] Nakoi Nitta, Feixiang Wu, Jung Tae Lee, and Gleb Yushin, *Li-ion battery materials: present and future*, Materials today (2015).

[4] Gregory L. Plett, *Equivalent-Circuit Methods*, Artech House, 2015.

[5] Roland Toth, *Modeling and identification of linear parameter-varying systems*, Springer, 2010.

[6] Roland Toth, Paul M J Van den Hof, and P.s.C Heuberger, *Model Structures for Identification of Linear Parameter-Varying (LPV) Models*, Proceedings of the Workshop on Systems and Control Theory in honor of József Bokor on his 60th birthday (2009), 15-34.

[7] Marian Gilson, Huges Garnier, and Roland Tóth, *Direct identification of continuous-time LPV models*, Proceedings of the American Control Conference (2011).

[8] Y. Hu, S. Yurkovich, Y. Guezennec, and B.J. Yurkovich, *A technique for dynamic battery model identification in automotive applications using linear parameter varying structures*, Control Engineering Practice **17** (2009), 1190-1201.

[9] Jürgen Remmlinger, Michael Buchholz, Thomas Soczka-Guth, and Klaus Dietmayer, *On-board state-of-health monitoring of lithium-ion batteries using linear parameter-varying models*, Journal of Power Sources **239** (2013), 689-695.

[10] Jürgen Remmlinger, Michael Buchholz, and Klaus Dietmayer, *Identification of a Bilinear and Parameter-varying Model for Lithium-ion Batteries by Subspace Methods*, American Control Conference (2013).

[11] Yiran Hu and Stephen Yurkovich, *Battery State of Charge Estimation in Automotive Applications using LPV Techniques*, American Control Conference (2010).

[12] Malin Andersson, Mikael Johansson, and Verena Löfqvist Klass, *A Continuous-Time LPV Model for Battery State-of-health Estimation Using Real Vehicle Data*, IEEE Conference on Control Technology and Applications (2020).

[13] Yiran Hu, Stephen Yurkovich, Yann Guezennec, and Raffaele Bornatico, *Model-Based Calibration for Battery Characterization in HEV Applications*, American Control Conference (2008).

[14] Andreas Cloppenborg, Carlos Rojas, Pablo Gonzalez, Joachim Horn, and Herbert Werner, *Nonlinear Model Predictive Control of a PEM Fuel Cell with an integrated Li-ion Battery using qLPV*, IEEE Conference on Control Technology and Applications (2019).

[15] Gregory L. Plett, *Battery modeling*, Artech House, 2015.

[16] Torsten Bohlin, *Practical Grey-box Process Identification*, Springer, 2006.

[17] European Commission, *Standards for the performance and durability assessment of electric vehicle batteries*, JRC Technical Reports (2018).

[18] Monowar Hossain, M.E Haque, Sajeeb Saha, and Mohammad Taufiqul Arif, *A Parameter Extraction Method for the Thevenin Equivalent Circuit Model of Liion Batteries*, IEEE Industry Applications Society Annual Meeting (2019).

[19] Rui Xiong, Hongwen He, Hongqiang Guo, and Yin Ding, *Modeling for Lithium-Ion Battery used in Electric Vehicles*, Procedia Engineering **15** (2011), 2869-2874.

[20] Ning Tian and Yebin Wang, *On parameter identification of an equivalent circuit model for lithium-ion batteries*, IEEE Conference on Control Technology and Applications (2017).

[21] Wladislaw Waag, Stefan Käbitz, and Dirk Uwe Sauer, *Experimental investigation of the lithium-ion battery impedance characteristic at various conditions and aging states and its influence on the application*, Applied Energy **102** (2013), 885-897.

[22] Mathworks, *System Identification Toolbox*, https://se.mathworks.com/products/sysid.html. Accessed February 19, 2020.

[23] _____, *Simulink Design Optimization Toolbox*, https://se.mathworks.com/products/sl-design-optimization.html. Accessed February 19, 2020.

[24] Ryan Ahmed, Javier Gazzari, Simona Onori, Saeid Habibi, Robyn Jackey, Kevin Rzemien, Jimi Tjong, and Jonathan LeSage, *Model-Based Parameter Identification of Healthy and Aged Li-ion Batteries for Electric Vehicle Applications*, SAE International (2015).

[25] Henry Miniguano, Andrés Barrado, Antonio Lázaro, Pablo Zumel, and Cristina Fernández, *General Parameter Identification Procedure and Comparative Study of Li-Ion Battery Models*, IEEE Transactions on Vehicular Technology (2020).

[26] Mohamed Daowd, Noshin Omar, Bavo Verbrugge, Peter Van Den Bossche, and Joeri Van Mierlo, *Battery Models Parameter Estimation based on Matlab/Simulink*, The 25th World Battery, Hybrid and Fuel Cell Electric Vehicle Symposium and Exhibition (2010).

[27] Mathworks, *Curve Fitting Toolbox*, https://se.mathworks.com/products/curvefitting.html. Accessed July 4, 2020.