CHALMERS
UNIVERSITY OF TECHNOLOGY

Ensemble Kalman Filter Design for Battery Management Systems

State and Parameter Estimation in Li-Ion Battery Cells

Master's thesis in Systems, Control and Mechatronics

LARS SJÖBERG
SANGEETHA NAGARAJU

Department of Electrical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2017

MASTER'S THESIS 2017:EX037

Ensemble Kalman Filter Design for Battery Management Systems

State and Parameter Estimation in Li-Ion Battery Cells

LARS SJÖBERG
SANGEETHA NAGARAJU

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Division of Signal Processing and Biomedical Engineering
Signal Processing Group
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2017

Ensemble Kalman Filter Design for Battery Management Systems
State and Parameter Estimation in Li-Ion Battery Cells
LARS SJÖBERG, SANGEETHA NAGARAJU

Supervisor: Dr. Hessam Mahdianfar, NEVS
Examiner: Dr. Karl Granström, Department of Electrical Engineering

Master's Thesis 2017:EX037
Department of Electrical Engineering
Division of Signal Processing and Biomedical Engineering
Signal Processing Group
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Concept for the NEVS 9-3 electric vehicle. Source: NEVS.

Typeset in LaTeX
Gothenburg, Sweden 2017

Ensemble Kalman Filter Design for Battery Management Systems
State and Parameter Estimation in Li-Ion Battery Cells
LARS SJÖBERG, SANGEETHA NAGARAJU
Department of Electrical Engineering
Chalmers University of Technology

Abstract

Increasing environmental awareness has in recent years motivated the automotive industry to advance towards electric and hybrid electric vehicles. Safe and efficient operation of electrically powered vehicles requires a wide range of systems and functions for managing power storage in battery packs. The performance of these systems is crucially dependent on knowledge about the remaining available charge in the batteries, referred to as state-of-charge (SOC). However, due to the complex electrochemical processes taking place inside battery cells, the SOC is generally hard to determine directly. As this is an important topic for further development of electrically powered vehicles, researchers are exploring the problem of efficiently and accurately estimating the battery cell SOC.

In this thesis, we investigate a novel model-based statistical filtering approach to the SOC estimation problem. In the Bayesian inference framework, an Ensemble Kalman Filter (EnKF) is implemented to jointly estimate inner battery states and parameters. For efficiency, the battery cell is modelled by a simple Equivalent Circuit Model (ECM), with three parameters that are initially identified using current and voltage measurements. With the objective to determine its viability for the proposed problem, the EnKF is compared to two previously studied and closely related algorithms: the Unscented Kalman Filter (UKF) and the bootstrap Particle Filter (PF).

The work has been carried out at National Electric Vehicle Sweden using pre-recorded data of current and voltage measurements for two different lithium-ion battery cell chemistries. Our findings suggest that the accuracy of the EnKF is similar to that of the PF, but with a lower computational cost, and overall more accurate than the UKF. In comparison, the EnKF is also easily implemented, with less tuning parameters than the UKF, and does not require as high dimensionality as the PF for convergence of the estimate. As such, we see good potential for the EnKF in the SOC estimation problem, providing flexibility and high performance.

Keywords: Electric vehicle, battery management system, state-of-charge, Bayesian filtering, ensemble Kalman filter, battery modelling, equivalent circuit model, unscented Kalman filter, particle filter.

Acknowledgements

Before introducing this thesis, we would like to provide this foreword expressing our most sincere appreciation and gratitude to a number of people that has been involved in the processes of creating, refining, and completing this project.

First, to National Electric Vehicles Sweden (NEVS) in Trollhättan, for providing us with the opportunity and the required resources to carry out this thesis. We would especially like to thank Dr. Hessam Mahdianfar, our supervisor at NEVS, for guiding us with his knowledge and experience in the field of advanced battery technologies throughout this thesis work.

Furthermore, to Dr. Karl Granström, our examiner at Chalmers University of Technology, for his valuable inputs and constant support. Additionally, we would like to thank everyone involved in the Master's thesis school for projects concerning estimation, filtering, tracking, mapping, localisation, and deep learning at Chalmers' signal processing group, for arranging and participating in valuable discussions and seminars.

Also, we would like to thank Alexander Berg and Andreas Käll, our peers at Chalmers, for helping us with improving our report by providing valuable and constructive feedback, as well as Deepak Kiran and Max Löfvall, our thesis opponents.

Special thanks also go to Dr. Gregory L. Plett, professor at the University of Colorado in Colorado Springs, for his vision of publicly accessible high-quality data; promoting and facilitating research in the field of advanced battery technologies.

Finally, we would like to thank and appreciate everyone who in one or the other way has helped us in our thesis work.

Lars Sjöberg and Sangeetha Nagaraju, Gothenburg, June 2017

Contents

List of Figures

Nomenclature

Notational Conventions

$\mathbf{A}^\mathsf{T}$	Transpose of matrix
$\mathrm{Cov}\left[\mathbf{x}\right]$	Covariance of $\mathbf{x}$
$\mathrm{E}\left[\mathbf{x}\right]$	Expectation of $\mathbf{x}$
$\mathrm{N}(\mathbf{x}; \mathbf{m}, \mathbf{P})$	The random variable $\mathbf{x}$ is Gaussian with mean $\mathbf{m}$ and covariance $\mathbf{P}$
$\sqrt{\mathbf{P}}$	Matrix such that $\mathbf{P} = \sqrt{\mathbf{P}}\sqrt{\mathbf{P}}^\mathsf{T}$
$p(\mathbf{x})$	Probability density of continuous or discrete random variable $\mathbf{x}$
$\mathbf{x}_{0:k}$	Sequence of vectors $\{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k\}$
$\bar{\mathbf{x}}$	Sample average along the second dimension of a set of vectors $\{\mathbf{x}^{(i)} : i = 1, \ldots, N\}$
$\dot{x}(t)$	Time derivative of the signal $x(t)$
$\hat{\mathbf{x}}$	Estimate of $\mathbf{x}$
$\mathbf{x} \sim p(\mathbf{x})$	The random variable $\mathbf{x}$ has the probability distribution $p(\mathbf{x})$

Symbols

$\mathbb{1}$	Vector of ones, $\mathbb{1} \triangleq [1, \ldots, 1]^\mathsf{T}$
$\mathbf{A}_k$	Dynamic model matrix in a linear time-varying model; transition matrix between time steps k and $k + 1$
α	Parameter of the unscented transform or stoichiometry related to state-of-charge
$A(q)$	Time shift polynomial of the system output
β	Parameter of the unscented transform
$\mathbf{B}_k$	Input matrix in a linear time-varying model at step k
$B(q)$	Time shift polynomial of the system input
C_1	Dynamic polarisation capacitance in equivalent circuit model
$\mathbf{C}_k$	Cross-covariance matrix of Kalman filter at time step k
$C(q)$	Moving average polynomial of the system disturbance
c_s	Concentration of lithium in the solid particles of a lithium-ion cell anode
$\delta(\cdot)$	The Dirac delta function
$D(q)$	Dynamics polynomial of the system disturbance

$\mathbf{E}_k$	Measurement noise ensemble matrix with columns $\mathbf{r}_k^{(i)}$ at time step k
ε	Estimation error, difference between true and estimated value $\varepsilon \triangleq x - \hat{x}$
η	Battery cell coulombic efficiency
$\mathbf{f}(\cdot)$	Dynamic transition function in a state-space model
$\mathbf{h}(\cdot)$	Measurement model function in a state-space model
$\mathbf{H}_k$	Measurement model matrix in a linear time-varying model at time step k
i	Battery cell load current or integer valued index variable
$\mathbf{I}_N$	Identity matrix of size N
k	Integer valued time step index
κ	Parameter of the unscented transform
$\mathbf{K}_k$	Gain matrix of Kalman filter at time step k
λ	Parameter of the unscented transform
m	Dimensionality of a measurement
N	Positive integer, in particular the number of samples in a Monte Carlo based filter
n	Dimensionality of the state
$\mathbf{P}$	Covariance of Gaussian distribution
p	Dimensionality of a system input
$\mathbf{P}_{k\|k-1}$	Predicted covariance of a Kalman filter at time step k, before the measurement $\mathbf{y}_k$
$\mathbf{P}_{k\|k}$	Covariance of a Kalman filter at time step k, after update by the measurement $\mathbf{y}_k$
$\mathbf{Q}$	Covariance matrix of process noise in a time-invariant model
Q	Battery cell total capacity
q	Time shift operator
$\mathbf{Q}_k$	Covariance matrix of process noise from time steps k to $k+1$
$\mathbf{q}_k$	Vector of Gaussian process noises
$\mathbf{R}$	Covariance matrix of measurement noise in a time-invariant model
R_0	Ohmic resistance in equivalent circuit model
R_1	Dynamic polarisation resistance in equivalent circuit model
$\mathbf{R}_k$	Covariance matrix of measurement noise at time step k
$\mathbf{r}_k$	Vector of Gaussian measurement noises
s	The complex Laplace transform variable
$\mathbf{S}_k$	Innovation covariance of Kalman filter at time step k
T	Battery cell temperature or index of the last time step, the final time of a time interval, $k \in [0, T]$
t	Continuous time variable, $t \in [0, \infty)$
$\boldsymbol{\theta}$	Vector of equivalent-circuit model parameters
$\boldsymbol{\theta}_k$	Vector of equivalent-circuit model parameters at time step k
$\boldsymbol{\vartheta}$	Vector of linear time-series model parameters
t_k	Time at step k such that $x(t_k) = x(kT_s)$
T_s	Sampling time interval
$\mathbf{u}_k$	Input vector at time step k
$u(t)$	System input signal as a function of time
v	Battery cell terminal voltage

v_h	Battery cell open-circuit voltage in a fully charged state
$\mathbf{V}_k$	Process noise ensemble matrix with columns $\mathbf{q}_k^{(i)}$ at time step k
v_l	Battery cell open-circuit voltage in a fully discharged state
v_{oc}	Battery cell open-circuit voltage
$W_i^{(c)}$	The ith covariance weight of the unscented transform
$W_i^{(m)}$	The ith mean weight of the unscented transform
$\mathbf{x}$	State vector or vector of random variables
$\mathcal{X}$	Sigma-point of the state $\mathbf{x}$
$\mathcal{X}_{k-1}$	Sigma-point of the state $\mathbf{x}_{k-1}$
$\mathcal{X}_k$	Sigma-point of the state $\mathbf{x}_k$
$\mathbf{x}^{(i)}$	The ith Monte Carlo draw from the distribution of $\mathbf{x}$
$\mathbf{X}_k$	State ensemble matrix with columns $\mathbf{x}_k^{(i)}$ at time step k
$\mathbf{x}_k$	State vector at time step k
$\hat{\mathbf{x}}_{k\|k-1}$	Predicted mean of a Kalman filter at time step k, before the measurement $\mathbf{y}_k$
$\hat{\mathbf{x}}_{k\|k}$	Mean of a Kalman filter at time step k, updated by the measurement $\mathbf{y}_k$
$\widetilde{\mathbf{X}}_k$	State ensemble matrix of deviations from the mean, $\mathbf{X}_k - \bar{\mathbf{x}}_k \mathbb{1}^\mathsf{T}$, at time step k
$\mathbf{y}$	Vector of measurements or random variables
$\mathbf{Y}_k$	Output ensemble matrix with columns $\mathbf{y}_k^{(i)}$ at time step k
$\mathbf{y}_k$	Vector of measurements at time step k
$\hat{\mathbf{y}}_{k\|k-1}$	Predicted mean of the measurement in a Kalman filter at time step k
$\widetilde{\mathbf{Y}}_k$	Output ensemble matrix of deviations from the mean, $\mathbf{Y}_k - \bar{\mathbf{y}}_k \mathbb{1}^\mathsf{T}$, at time step k
z	Battery cell state-of-charge or the complex $\mathscr{Z}$-transform variable

Acronyms

AEKF	Adaptive Extended Kalman Filter
ARMAX	Auto Regressive Moving Average with eXogenous input
ARX	Auto Regressive with eXogenous input
BJ	Box-Jenkins
BMS	Battery Management System
CDKF	Central-Difference Kalman Filter
ECM	Equivalent Circuit Model
EKF	Extended Kalman Filter
EnKF	Ensemble Kalman Filter
EV	Electric Vehicle
HEV	Hybrid Electric Vehicle
KF	Kalman Filter
KVL	Kirchhoff's Voltage Law
LMMSE	Linear Minimum Mean Square Error
LUT	Look-Up Table
MAE	Mean Absolute Error
MAP	Maximum A Posteriori

OCV	Open-Circuit Voltage
OE	Output-Error
PF	Particle Filter
RC	Resistor–Capacitor
RMSE	Root-Mean-Square Error
SOC	State Of Charge
STD	STandard Deviation
UDDS	Urban Dynamometer Drive Schedule
UKF	Unscented Kalman Filter
UT	Unscented Transform

1
Introduction

Vehicles powered by electric motors have many advantages over the traditional choice of internal combustion engines. Environmental impact, energy efficiency, mechanical complexity, and noise pollution are only a few. Safe and efficient operation of Electric Vehicles (EVs) is crucially dependent on proper management of the on-board power storage; most commonly in lithium-ion battery packs. The embedded system employed in an electrically powered application for sensing, protection, interfacing, managing performance, and diagnosing of battery packs is collectively labelled as a Battery Management System (BMS). One of the key factors for reliable performance management in a BMS is knowledge about the individual battery cells' State Of Charge (SOC)—the available electrical capacity at any given time. The average SOC of a complete battery pack can thus, in essence, be compared to the fuel gauge in an internal combustion engine vehicle.

Electrochemically, the cell SOC is a stoichiometry, related to the average concentration of lithium in the negative electrode solid particles [1], and can as such in practice not be measured directly. This motivates the employment of more advanced techniques for estimating the SOC, which is typically challenged by noise-contaminated sensor measurements, variations in temperature, and battery cell degeneracy [2] as well as an overall complex, nonlinear, and time-variant dynamic behaviour. Methods are thus needed for combining information from several sources to produce a less uncertain estimate.

The problem of estimating inner, and oftentimes immeasurable, states of a system or process is an important topic in signal processing applications. Advanced estimation of SOC in battery cells is commonly approached using statistical *filtering* techniques; specifically in the framework of the Kalman Filter (KF) [3] and its relatives. The objective of this thesis is to investigate a novel KF based approach to the SOC estimation problem—the so-called Ensemble Kalman Filter (EnKF)—in an attempt to fill the gap between the commonly exploited methods.

1.1 Background

State estimation in electrochemical battery cells is a relatively new research topic, but has been investigated thoroughly during the last few decades, much motivated by the resurgence of EVs and Hybrid Electric Vehicles (HEVs) in the 1990s [4], in which power is managed and stored in battery packs of multiple lithium-ion cells. In particular, the battery cell SOC is of high importance for energy and power calculations in EV applications [5], and serves as a crucial parameter in, e.g., determining an efficient control strategy, ensuring safe power utilisation, and improving the battery life [6].

For the purpose of real-time SOC estimation, a wide range of techniques have been

studied [7], varying in complexity, accuracy, and versatility. Real-time estimation of SOC is of particular interest in the automotive sector and demands indirect methods that are fast, robust, and accurate. The most popular methods are, in brief [5]:

- *Current-based methods:* Exploits the fact that the supplied or withdrawn current is directly dependent on the change of available charge in the battery; often referred to as "coulomb counting". Current measurements are continuously used to update the SOC estimate.
- *Voltage-based methods:* Offline determination of the relationship between the SOC and the battery cell's Open-Circuit Voltage (OCV) enables online look-up of SOC, given terminal voltage measurements and a mathematical battery model.
- *Model-based state estimation:* Employs mathematical models to make predictions for inner process states as well as impending measurements. When the real system is observed, associated predicted and measured quantities are compared, and the potential difference is fed back to update the predicted inner states. This scheme is illustrated in Figure 1.1.

The model-based state estimation technique provides a number of advantages—much due to the information provided by the feedback of the prediction error—and is often the preferred choice when computational capacity is sufficient.

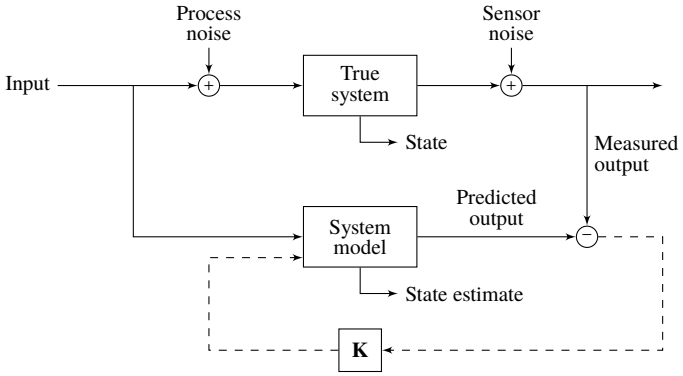


Figure 1.1: Conceptual schematic of the model-based approach for state estimation. The difference between predicted and measured system output is fed back, under some policy **K**, to the system model to update the state estimate.

Under a certain set of conditions one such model-based state estimator, the KF, computes the provably *optimal* state estimate in real-time by combining modelled and observed data. Optimality is obtained in the minimum-variance sense, and is limited to linear systems, by proper selection of the feedback policy **K**, cf. Figure 1.1. The estimate is delivered together with a dynamic uncertainty of the approximation, weighing the user-specified quality of the model prediction against the quality of the measurements, and making the Kalman filter an exceptionally useful algorithm for the problem at hand.

One of the key challenges encountered in SOC estimation is the nonlinear mapping between observed data and inner battery states. This demands usage of suboptimal derivatives of the KF, as the linear assumption is not fulfilled. Another challenge in this problem is the inherent dependence of the electrochemical dynamics with respect to the SOC; such that the battery model dynamics change nonlinearly during cell charge and

discharge. This is usually approached by also estimating the model parameters as SOC changes.

1.2 Problem Statement

In this thesis, a statistical inverse problem is investigated for estimating inner states and parameters in lithium-ion battery cells within the inference framework of Bayesian probability theory. In this context, the estimate of a target quantity $\mathbf{x}$ is, in essence, computed as a conditional probability distribution of the quantity, given available measurements $\mathbf{y}$, by application of optimal filtering theory. That is: for every measurement of the process an estimate of the SOC is computed together with an associated uncertainty.

Conceptually, the proposed problem can be divided into the following four subtasks.

1. **Dynamical model.** Mathematical modelling of the electrochemical dynamics of lithium-ion battery cells. In the Bayesian inference framework, prior knowledge about the state $\mathbf{x}$ is represented by a probabilistic dynamical model, describing the transition of battery states and parameters in time.

2. **Measurement model.** Mathematical modelling of the sensor measurements $\mathbf{y}$, as related to the process states $\mathbf{x}$. This probabilistic measurement model represents the likelihood of the measurements, given the prior knowledge of the state.

3. **Filter design.** Design and implementation of the estimation scheme for fusing the information provided by the dynamical and measurement models with measurement data gained from observations. This fusion process is referred to as *filtering*.

4. **Evaluation.** Comparative evaluation of the developed filtering algorithm as related to similar approaches, represented in external research.

Task 3 constitutes the main objective of this thesis and represents the motivation for our investigation: to explore the literature gap in SOC estimation, between the two more common and closely related sample-based filtering techniques referred to as *sigma-point methods* and *particle filters*.

1.3 Purpose and Thesis Objective

The purpose of this thesis is to design, implement, and evaluate an EnKF for real-time estimation of SOC in lithium-ion battery cells. The battery model used to predict the evolution of the system is adapted by also estimating associated parameters. The performance of the EnKF algorithm shall be assessed with regard to closely related filtering algorithms—the Unscented Kalman Filter (UKF) and the bootstrap Particle Filter (PF).

The main objective of this thesis is thus to investigate the EnKF as an alternative nonlinear filtering algorithm in SOC estimation. The following research questions highlight this aim within the general purpose of the thesis.

(i) What is a suitable battery model for prediction in real-time SOC estimation, with regard to accuracy and computational complexity?

(ii) Can the battery model parameters be estimated on-line to increase the accuracy of the SOC estimate?
(iii) What is the potential of the EnKF for the proposed SOC estimation problem and how does it compare to the PF and UKF?
(iv) With regard to tunable filter parameters; what are the minimum criteria of the EnKF that ensures filter converges?

1.4 Scope and Limitations

In this thesis, the estimate of the target quantity $\mathbf{x}$ is considered to be a Gaussian random variable. That is, the result of the estimation scheme is characterised by the mean and covariance of a Gaussian distribution, representing the actual estimate and its associated uncertainty.

It is generally required to perform time-consuming and expensive laboratory tests, in order to obtain accurate offline estimates of SOC and the nonlinear mapping to the cell OCV. This thesis does not aim to produce such data but shall use readily available data sets provided by Dr. Gregory L. Plett at the University of Colorado, Colorado Springs, (UCCS). The details of these resources are explained in Section 4.3.

The posed problem in Section 1.2 shall be solved within the frames of applicability to EVs and thus expected power demand during driving. The models for predicting the battery cell state evolution are therefore required to be simple enough for implementation on embedded computer systems. This requirement is validated by reviewing the literature on the topic of battery modelling, and will not be subject to actual tests on embedded hardware. As such, all algorithms are developed and tested in the MATLAB and SIMULINK software.

The developed estimation algorithm requires measurements of cell terminal voltage and load current, as well as information about the cell temperature. This information is in a real system acquired using sensors with an associated uncertainty, but due to restricted access to appropriate data sets the current will be regarded as a known input to the system. Furthermore, the battery cell temperature is assumed to be constant, uniformly distributed, and completely known throughout each test scenario. Note that inferring the battery cell's inner temperature can be regarded as an independent estimation problem, which is not of interest in this thesis.

In reality, battery cells suffer from degeneracy as the number of charge–discharge cycles increase. This has an effect on e.g. how much energy the cell can store and its charging efficiency. The time frame of these changes is however outside the scope of this project, and the effects associated with battery ageing are assumed to be negligible in the tests performed.

1.5 Related Work

The most distinct subproblems that arise in the topic of battery SOC estimation—modelling of battery dynamics and selection of a suitable estimation scheme—are both well-represented in the literature. Battery modelling is an important subject in both understanding the short and long-term behaviour of electrochemical power storage systems and improving

the quality of model-based estimators in EV and HEV applications. Advanced estimation schemes are needed to further develop the flexibility, performance, and robustness of electric vehicles, by providing more accurate information about the power state of the vehicle. Below follows a brief review of some of the most prominent contributions and related studies in these areas, as well as an introduction to the EnKF.

1.5.1 Battery Modelling

In particular, modelling of battery dynamics is a diverse and thoroughly investigated topic. The two most important categories for the wide range of existing approaches are *electrochemical* and *equivalent circuit* modelling.

Electrochemical models or *physics-based models* aim to represent the internal electrochemical processes occurring in the battery cell. These models are more detailed and computationally heavy, typically describing the cell dynamics using coupled partial differential equations [8].

Equivalent circuit models describe input–output behaviour of the battery cell. Empirical data is used to fit black-box model parameters, represented by electrical components. These low-complexity, and sometimes even linear, models generally yield robust and fast output predictions, but cannot provide any information about inner battery states [5].

Several review studies have been published for EV and HEV simulation [9] and real-time estimation of battery states [8], [10], [11], in detail discussing the two categories of modelling techniques above. Two recent studies [12], [13] in electrochemical lithium-ion battery modelling incorporate simplified physical models for real-time simulation and estimation of battery states, such as state-of-charge and state of health. Another recent example [14] presents a reduced-order electrochemical model for a composite electrode lithium-ion battery, a technology which has seen an increase in popularity for automotive applications. Furthermore, in [15], an electrochemical-polarisation model is proposed for real-time estimation, combining the electrochemical Nernst model with an Equivalent Circuit Model (ECM) for battery polarisation effects.

While the general conclusion in the literature is that the choice of model should reflect the intended application, the low-complexity ECMs are the best choice for real-time model-based estimation applications. In particular, the first and second order Resistor–Capacitor (RC) models, with extensions, have proven to be the preferred choice in SOC estimation problems [10].

1.5.2 Model-Based Estimation Techniques

As the practical standard choice for nonlinear filtering problems [16], the Extended Kalman Filter (EKF) has been investigated extensively for state estimation in battery management systems. In a series of three highly cited papers by G. L. Plett [17]–[19], the EKF algorithm is used with an ECM to estimate SOC and time-variant model parameters. Innovation-based adaptive versions of the EKF—often referred to as Adaptive Extended Kalman Filter (AEKF) in the literature—have shown to further improve SOC estimates [20], [21].

An accurate estimation of SOC depends on its parameters that vary along with it. This surfaces the necessity to jointly estimate the parameters along with the states of interest. In [22] an AEKF has been proposed for estimating the SOC along with the battery model parameters that vary with it; to account for the varying parameters the noise terms in the motion and measurement models are adapted based on innovation sequences inside a moving estimation window. A similar AEKF method proposed in [20] shows that the maximum SOC estimation error decreases from 14.96 % to 2.54 % and that the mean SOC estimation error reduces from 3.19 % to 1.06 %. A further more advanced method for jointly estimating SOC and peak power capabilities in lithium-ion batteries has been proposed in [23]. In this method, the joint estimation of both the states along with the varying parameters has been done by implementing an AEKF as well, but the information from the coupled states of peak power capability and SOC has facilitated to achieve an accurate SOC estimate with an estimation error below 2 %.

As the EKF comes with a palpable computational complexity, requiring linearisation of the dynamics and measurements models in every iteration of the algorithm, the more versatile sigma-point methods have been investigated as an alternative, producing overall better results than the EKF at the same computational cost. In [24], [25], Plett revisits his previous study and achieves better results with a Central-Difference Kalman Filter (CDKF) than with the previous EKF approach. The UKF is another common nonlinear filtering algorithm in the category of sigma-point methods [24], and has as well been studied for SOC estimation, by e.g. [26]–[28].

Less cited studies [29]–[31] have been performed using various versions of the highly flexible and computationally heavy PF. Results show more accurate estimation results for the PF, as compared to EKF and UKF algorithms, but at a higher computational cost.

1.5.3 The Ensemble Kalman Filter

The EnKF is, like the sigma-point methods and particle filters, a sample-based method for nonlinear filtering problems. Its ability to handle extremely high-dimensional, nonlinear, and non-Gaussian problems has made the EnKF a popular choice for estimation problems in geoscientific disciplines [32]. The algorithm was first proposed in 1994 by Geir Evensen [33]—and modified in [34] to account for a systematically underestimated uncertainty estimate—as a better alternative to nonlinear filtering problems than the EKF, specifically for ocean models.

Interestingly, the EnKF stands somewhat isolated in relation to recent signal processing literature, and with few references to comparable nonlinear filtering algorithms [32], despite similarities with sigma-point methods and particle filters. The majority of EnKF studies have been published within the geoscientific area, predominantly with applications in weather and ocean forecasting, under the terminology “data assimilation”. A few exceptions exist in the form of review articles [35]–[38].

1.6 Major Contributions

The scientific value of this thesis is centred around the investigation of the EnKF as a novel estimation algorithm in the SOC estimation problem. As highlighted by the related work, the EKF is a popular choice for the proposed problem and comes in many flavours and

configurations. The EKF uses a first order Taylor expansion to approximate nonlinearities in the problem, and thus requires the mathematical models of the process the sensors to be differentiable. For this reason, the EKF also suffers from a relatively high computational complexity and divergence of the estimate if the nonlinearities are considerable [39]. It is desirable to investigate if the EnKF can circumvent the issues associated with the EKF.

Other methods, such as the UKF, use sigma-points to better capture the effect of nonlinear transforms on the filtering distribution. However, the information contained in the sigma-points are in every iteration condensed into an estimate and an associated uncertainty, resulting in loss of information. The EnKF incorporates a set of samples, similar to the sigma-points, but keeps the information in the set of samples throughout the entire estimation sequence. It is of interest to analyse the effects of this property in the proposed problem, which is one of the contributions of this thesis.

The particle filter, being a versatile and computationally heavy filtering algorithm, is closely related to the EnKF. In this thesis, the novel EnKF approach provides similar versatility but at a lower computational cost.

1.7 Thesis Outline

The remainder of this thesis report is structured as follows. Chapter 2 presents the theoretical foundation required for performing statistical filtering. After a general review of the Bayesian inference framework, algorithms are formulated and discussed for the KF, the UKF, the bootstrap PF and the EnKF. Next, Chapter 3 describes the terminology and philosophy of equivalent circuit battery modelling, explaining the different concepts considered in the development of the dynamical model.

In Chapter 4, the estimation problem is formally defined followed by the system analysis required for identifying battery model parameters. The second half of Chapter 4 describes the test scenario, data, and performance metrics used for evaluating the different estimation schemes. Results are presented in Chapter 5 and discussed in Chapter 6. Finally, the thesis conclusions are given in Chapter 7.

2
Statistical Filtering

The objective of this thesis is to estimate a set of immeasurable states and parameters in a dynamical system. In this chapter, the topic of optimal filtering is briefly reviewed, as a model-based estimation approach. The chapter begins with a general description of the Bayesian statistical inference framework; stating important assumptions made regarding dynamical and stochastic processes. To support subsequent discussions about comparing different filtering algorithms, we present the linear KF, the UKF, the bootstrap PF, and finally the EnKF. In cases where a nonlinear filtering problem is considered, the estimate is approximated as a Gaussian density.

2.1 Bayesian Estimation

Optimal filtering refers to the problem of finding the best estimate of the inner state in a time-varying system, given indirect and noise-contaminated observations. In particular, the *filtering* problem concerns estimation of the unknown *present* state of the process given past and current observations. Mathematically, this is conventionally considered a statistical inversion problem, which in the *Bayesian* sense entails combining *prior* information about the process with present observations or measurements. The estimate is computed as a joint probability distribution of all the hidden states up until time k, $\mathbf{x}_{0:k}$, conditioned on all available observations, $\mathbf{y}_{1:k}$, by application of Bayes' rule,

$$p(\mathbf{x}_{0:k} \mid \mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_{1:k} \mid \mathbf{x}_{0:k})p(\mathbf{x}_{0:k})}{p(\mathbf{y}_{1:k})}, \tag{2.1}$$

describing the sought posterior distribution $p(\mathbf{x}_{0:k} \mid \mathbf{y}_{1:k})$ of the hidden states in terms of their prior distribution $p(\mathbf{x}_{0:k})$ and the measurement likelihood $p(\mathbf{y}_{1:k} \mid \mathbf{x}_{0:k})$. In this context, optimality has a statistical meaning signifying that the posterior estimate is the one with minimal uncertainty.

The state time sequence $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots\}$ is assumed to have the Markov property, which means that the present state of the process $\mathbf{x}_k$ is dependent only on the previous state $\mathbf{x}_{k-1}$; and independent of all other past states and observations,

$$p(\mathbf{x}_k \mid \mathbf{x}_{0:k-1}, \mathbf{y}_{1:k-1}) = p(\mathbf{x}_k \mid \mathbf{x}_{k-1}). \tag{2.2}$$

Consequently, any measurement in the sequence of observations $\{\mathbf{y}_1, \mathbf{y}_2, \dots\}$ is conditionally independent of all other measurements and only dependent on the current process state;

$$p(\mathbf{y}_k \mid \mathbf{x}_{0:k}, \mathbf{y}_{1:k-1}) = p(\mathbf{y}_k \mid \mathbf{x}_k). \tag{2.3}$$

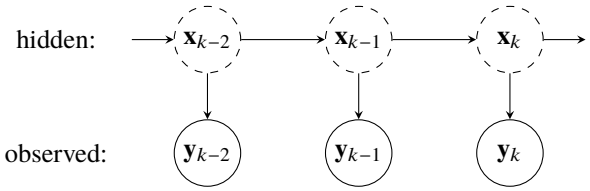


Figure 2.1: Bayesian network illustrating the conditional dependencies between the hidden state sequence $\mathbf{x}$ and observations $\mathbf{y}$.

These dependencies are illustrated in the form of a Bayesian network in Figure 2.1. Under the Markov assumption, the probability distributions in equation (2.1) can as such be described *recursively* by a probabilistic state-space model:

$$\mathbf{x}_0 \sim p(\mathbf{x}_0), \tag{2.4a}$$
$$\mathbf{x}_k \sim p(\mathbf{x}_k \mid \mathbf{x}_{k-1}), \tag{2.4b}$$
$$\mathbf{y}_k \sim p(\mathbf{y}_k \mid \mathbf{x}_k), \tag{2.4c}$$

specifying the initial state distribution of the process (2.4a), the state transition probability distribution (2.4b), and the dependence of observations on the hidden state, (2.4c).

The objective of the recursive Bayesian filter is to, at each time instant k, compute the marginal distribution of the present process state $\mathbf{x}_k$ given all available observations $\mathbf{y}_{1:k}$, i.e., $p(\mathbf{x}_k \mid \mathbf{y}_{1:k})$. Given a prior distribution, each subsequent filtering iteration consists of a *prediction* step and an *update* step. The predicted density at time k can, given the dynamical model (2.4b) and all previous measurements $\mathbf{y}_{1:k-1}$, be computed using the Chapman–Kolmogorov equation:

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} \mid \mathbf{y}_{1:k-1}) \, \mathrm{d}\mathbf{x}_{k-1}. \tag{2.5}$$

The posterior density at time k is then, given the measurement model (2.4c) and the current measurement $\mathbf{y}_k$, obtained by using Bayes' rule [39],

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_k \mid \mathbf{x}_k) p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1})}{p(\mathbf{y}_k \mid \mathbf{y}_{1:k-1})}. \tag{2.6}$$

The recursive Bayesian estimation scheme is illustrated in Figure 2.2 in terms of the densities in equations (2.5) and (2.6).

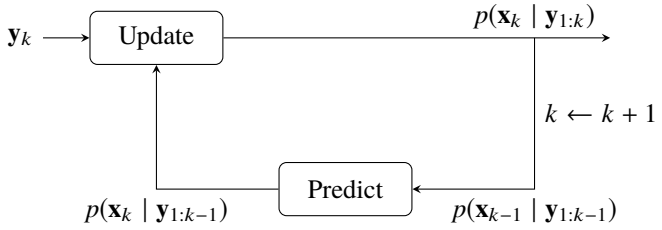


Figure 2.2: Recursive Bayesian estimation scheme. The predicted density $p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1})$ is computed using the *transition* density $p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$ and the previous posterior density $p(\mathbf{x}_{k-1} \mid \mathbf{y}_{1:k-1})$, after which measurements $\mathbf{y}_k$ are fused with the prediction to obtain the posterior density $p(\mathbf{x}_k \mid \mathbf{y}_{1:k})$.

2.2 The Kalman Filter

The Kalman filter [3] was first presented in 1960 by Rudolph E. Kálmán as a solution to the optimal filtering problem. The algorithm is exactly the closed form solution to the optimal filtering equations (2.5) and (2.6) when the probabilistic state-space model (2.4) in linear and Gaussian. In such a scenario, a matrix $\mathbf{A}_{k-1}$ can be used to describe the state transition from time step $k-1$ to k and another matrix $\mathbf{H}_k$ describes the linear mapping between the process states $\mathbf{x}_k \in \mathbb{R}^n$ and the observations $\mathbf{y}_k \in \mathbb{R}^m$. The filtering model can hereby be written as a set of discrete-time difference equations,

$$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{q}_{k-1}, \tag{2.7a}$$
$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{r}_k, \tag{2.7b}$$

where $\mathbf{q}_{k-1} \sim \mathrm{N}(\mathbf{0}, \mathbf{Q}_{k-1})$ and $\mathbf{r}_k \sim \mathrm{N}(\mathbf{0}, \mathbf{R}_k)$ are additive zero-mean Gaussian processes, with covariance matrices $\mathbf{Q}_{k-1}$ and $\mathbf{R}_k$ respectively, describing the stochastic part of the dynamical and measurement models.

In this setting, all densities involved in the optimal filtering equations (2.5) and (2.6) are guaranteed to be Gaussian [39], under the condition that the prior $p(\mathbf{x}_0)$ is Gaussian, and can thus be completely characterised by their first and second moments: the mean and the covariance. Denoting by $\hat{\mathbf{x}}$ the mean of the filtering distribution and by $\mathbf{P}$ the covariance, the KF can be summarised by Algorithm 1 below.

Algorithm 1 (Kalman filter)**.** *Given a linear and Gaussian filtering model* (2.7)*, the characterising moments of the optimal filtering density, in the variance-minimising sense, are computed by the following prediction and update steps:*

Prediction:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1|k-1}, \tag{2.8a}$$
$$\mathbf{P}_{k|k-1} = \mathbf{A}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{A}_{k-1}^{\mathsf{T}} + \mathbf{Q}_{k-1}. \tag{2.8b}$$

Update:

$$\mathbf{v}_k = \mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1}, \tag{2.9a}$$
$$\mathbf{S}_k = \mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^{\mathsf{T}} + \mathbf{R}_k, \tag{2.9b}$$
$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^{\mathsf{T}}\mathbf{S}_k^{-1}, \tag{2.9c}$$
$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\mathbf{v}_k, \tag{2.9d}$$
$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^{\mathsf{T}}. \tag{2.9e}$$

The densities of particular interest in the KF are the *predicted* distribution,

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1}) = \mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}),$$

the measurement prediction or *innovation* distribution,

$$p(\mathbf{y}_k \mid \mathbf{y}_{1:k-1}) = \mathrm{N}(\mathbf{y}_k; \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1}, \mathbf{S}_k)$$

and the *filtered* distribution,

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:k}) = \mathrm{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}). \tag{2.10}$$

This summary of the Kalman filter is concluded by making a final remark about the posterior distribution computed by (2.9). It can be shown that estimate $\hat{\mathbf{x}}_{k|k}$ is—by selection of the Kalman gain $\mathbf{K}_k$—optimal in the sense that it is the most probable value of the random variable $\mathbf{x}$ at time k, as well as the value that minimises the mean squared error of the estimate [40]. That is, the Kalman filter computes both the *maximum a posteriori* (MAP) and *(linear)*[1] *minimum mean square error* (LMMSE) estimate of the state $\mathbf{x}$.

The following sections will present Kalman filter inspired techniques for solving *nonlinear* filtering problems, in which approximations have to be made in order to not deviate from the Gaussian framework.

2.3 The Unscented Kalman Filter

The UKF is a popular filtering algorithm for nonlinear estimation problems. It was proposed by Julier and Uhlmann in 1997 [41] as a better alternative to the EKF, which commonly suffers from being generally difficult to implement and tune. The UKF belongs to a class of filtering techniques called sigma-point methods, which employ a set of discretely sampled "sigma-points" in the state-space to approximate mean and covariance of the filtering distribution as it is subjected to the possibly nonlinear transforms present in the motion and measurement models.

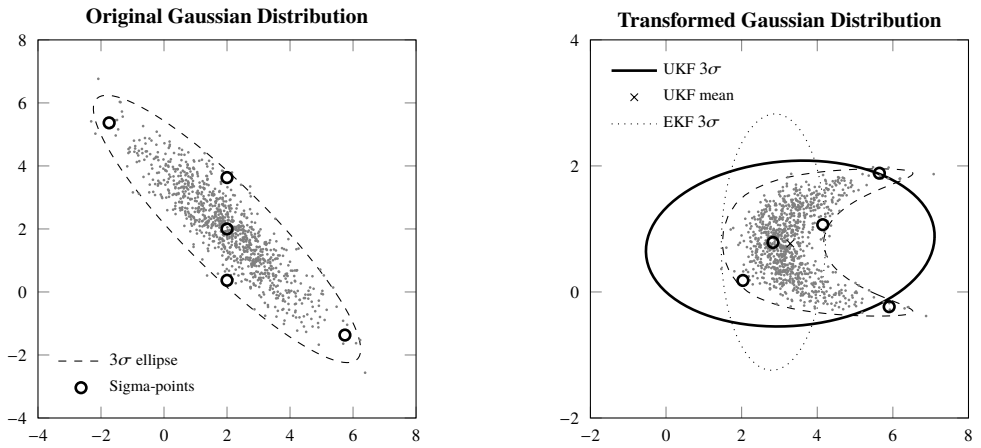


Figure 2.3: Illustration of the unscented transform based approximation of a transformed two-dimensional Gaussian random variable. To the left is the distribution of the original Gaussian random variable, illustrated with samples and the contour of the three standard deviation (3σ) ellipse (dashed). A total of 5 sigma-points are placed to capture the original distribution, using scaling parameters $\alpha = 1, \beta = 2, \kappa = 5$. To the right is the transformed distribution illustrated together with the transformed sigma-points. The unscented transform Gaussian approximation is drawn with a solid thick line, representing the 3σ ellipse, and a cross representing the approximated mean. Also drawn is the linear approximation that an EKF would produce (dotted).

In contrast to the linear filtering model in (2.7), the UKF is formulated for a nonlinear

[1]The Kalman filter is restricted to be an affine approximation of data.

stochastic state-space model of the form

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{q}_{k-1}, \tag{2.11a}$$
$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{r}_k, \tag{2.11b}$$

where $\mathbf{f} : \mathbb{R}^n \mapsto \mathbb{R}^n$ and $\mathbf{h} : \mathbb{R}^n \mapsto \mathbb{R}^m$ are the nonlinear state transition and measurement functions; and $\mathbf{q}$ and $\mathbf{r}$ are as in (2.7). Since $\mathbf{f}$ and $\mathbf{h}$ are assumed to be nonlinear, the transformed random variable, $\mathbf{f}(\mathbf{x})$ or $\mathbf{h}(\mathbf{x})$, does not necessarily retain its Gaussian property. The UKF forms a Gaussian approximation to the transformed density by computing the *weighted* average of the transformed sigma-points, serving as an estimate for the mean, and their *weighted* outer product, estimating the covariance [42].

The Gaussian approximation of the filtering distribution employed in the UKF is based on a method called the Unscented Transform (UT), which is also credited to Julier and Uhlmann. For a general Gaussian random variable $\mathbf{x}$, with mean $\mathbf{m}$ and covariance $\mathbf{P}$, and nonlinear transform $\mathbf{g}$,

$$\begin{aligned} \mathbf{x} &\sim \mathrm{N}(\mathbf{m}, \mathbf{P}), \\ \mathbf{y} &= \mathbf{g}(\mathbf{x}), \end{aligned} \tag{2.12}$$

the UT computes the Gaussian approximation of the transformed random variable $\mathbf{y}$ by first forming a set of $2n + 1$ sigma-points as

$$\begin{aligned} \mathcal{X}^{(0)} &= \mathbf{m}, \\ \mathcal{X}^{(i)} &= \mathbf{m} + \sqrt{n + \lambda} \left[\sqrt{\mathbf{P}}\right]_i, \quad i = 1, \ldots n, \\ \mathcal{X}^{(i+n)} &= \mathbf{m} - \sqrt{n + \lambda} \left[\sqrt{\mathbf{P}}\right]_i, \quad i = 1, \ldots n, \end{aligned} \tag{2.13}$$

where λ is a scaling parameter defined as

$$\lambda = \alpha^2(n + \kappa) - n. \tag{2.14}$$

The parameters α and κ can be adjusted to control the spread of the sigma-points; see example in Figure 2.3. The notation $[\sqrt{\mathbf{P}}]_i$ in (2.13) denotes the ith column of a matrix $\sqrt{\mathbf{P}}$ such that $\sqrt{\mathbf{P}}\sqrt{\mathbf{P}}^{\mathsf{T}} = \mathbf{P}$.[2] By propagating the sigma-points through the nonlinear transform,

$$\mathcal{Y}^{(i)} = \mathbf{g}\big(\mathcal{X}^{(i)}\big), \quad i = 0, \ldots, 2n, \tag{2.15}$$

the mean and covariance of the transformed random variable are approximated by

$$\begin{aligned} \mathrm{E}\left[\mathbf{g}(\mathbf{x})\right] = \bar{\mathbf{y}} &\approx \sum_{i=0}^{2n} W_i^{(m)} \mathcal{Y}^{(i)}, \\ \mathrm{Cov}\left[\mathbf{g}(\mathbf{x})\right] = \bar{\mathbf{P}}_{\mathbf{y}} &\approx \sum_{i=0}^{2n} W_i^{(c)} \left(\mathcal{Y}^{(i)} - \bar{\mathbf{y}}\right)\left(\mathcal{Y}^{(i)} - \bar{\mathbf{y}}\right)^{\mathsf{T}}, \end{aligned} \tag{2.16}$$

[2] There are a number of methods available for computing the square root of the state estimate covariance matrix. For instance, using eigenvalue decomposition, Shur decomposition, or Cholesky decomposition. The latter method is indicated to be best suited for the UKF algorithm [43].

using the weights $W_i^{(m)}$ and $W_i^{(c)}$ defined as

$$\begin{aligned} W_0^{(m)} &= \frac{\lambda}{n+\lambda}, \\ W_0^{(c)} &= \frac{\lambda}{n+\lambda} + (1-\alpha^2+\beta), \\ W_i^{(m)} &= \frac{1}{2(n+\lambda)}, \quad i = 1, \dots 2n, \\ W_i^{(c)} &= \frac{1}{2(n+\lambda)}, \quad i = 1, \dots 2n. \end{aligned} \tag{2.17}$$

For Gaussian distributions, a value of $\beta = 2$ is optimal in the calculation (2.17) of the UT weights. An example of the UT Gaussian approximation method is illustrated in Figure 2.3, for a two-dimensional Gaussian random variable subjected to a nonlinear transform.

Using equations (2.13)–(2.17), the UKF algorithm has the formulation given in Algorithm 2 [39].

Algorithm 2 (Unscented Kalman Filter)**.** *The unscented Kalman filter for nonlinear problems of the form* (2.11) *computes the filtering distribution at each time step k in the following prediction and update steps:*

Prediction: *Form the sigma-points*

$$\begin{aligned} \mathcal{X}_{k-1}^{(0)} &= \hat{\mathbf{x}}_{k-1|k-1}, \\ \mathcal{X}_{k-1}^{(i)} &= \hat{\mathbf{x}}_{k-1|k-1} + \sqrt{n+\lambda}\left[\sqrt{\mathbf{P}_{k-1|k-1}}\right]_i, \quad i = 1, \dots n, \\ \mathcal{X}_{k-1}^{(i+n)} &= \hat{\mathbf{x}}_{k-1|k-1} - \sqrt{n+\lambda}\left[\sqrt{\mathbf{P}_{k-1|k-1}}\right]_i, \quad i = 1, \dots n, \end{aligned} \tag{2.18}$$

and compute the predicted mean and covariance after propagating the sigma-points through the dynamic model:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= \sum_{i=0}^{2n} W_i^{(m)} \mathbf{f}(\mathcal{X}_{k-1}^{(i)}), \\ \mathbf{P}_{k|k-1} &= \sum_{i=0}^{2n} W_i^{(c)} \left(\mathbf{f}(\mathcal{X}_{k-1}^{(i)}) - \hat{\mathbf{x}}_{k|k-1}\right)\left(\mathbf{f}(\mathcal{X}_{k-1}^{(i)}) - \hat{\mathbf{x}}_{k|k-1}\right)^{\mathsf{T}}, \end{aligned} \tag{2.19}$$

using the weights $W_i^{(m)}$ and $W_i^{(c)}$ defined in equation (2.17)*.*

Update: *Form the sigma-points*

$$\begin{aligned} \mathcal{X}_{k}^{(0)} &= \hat{\mathbf{x}}_{k|k-1}, \\ \mathcal{X}_{k}^{(i)} &= \hat{\mathbf{x}}_{k|k-1} + \sqrt{n+\lambda}\left[\sqrt{\mathbf{P}_{k|k-1}}\right]_i, \quad i = 1, \dots n, \\ \mathcal{X}_{k}^{(i+n)} &= \hat{\mathbf{x}}_{k|k-1} - \sqrt{n+\lambda}\left[\sqrt{\mathbf{P}_{k|k-1}}\right]_i, \quad i = 1, \dots n, \end{aligned} \tag{2.20}$$

and compute the mean and covariance of the predicted output, as well as the cross-covariance of the state and the measurement, after propagating the sigma-points

through the measurement model:

$$\begin{aligned}\hat{\mathbf{y}}_{k|k-1} &= \sum_{i=0}^{2n} W_i^{(m)} \mathbf{h}(\mathcal{X}_k^{(i)}),\\ \mathbf{S}_k &= \sum_{i=0}^{2n} W_i^{(c)} \left(\mathbf{h}(\mathcal{X}_k^{(i)}) - \hat{\mathbf{y}}_{k|k-1}\right)\left(\mathbf{h}(\mathcal{X}_k^{(i)}) - \hat{\mathbf{y}}_{k|k-1}\right)^{\mathsf{T}},\\ \mathbf{C}_k &= \sum_{i=0}^{2n} W_i^{(c)} \left(\mathcal{X}_k^{(i)} - \hat{\mathbf{x}}_{k|k-1}\right)\left(\mathbf{h}(\mathcal{X}_k^{(i)}) - \hat{\mathbf{y}}_{k|k-1}\right)^{\mathsf{T}}.\end{aligned} \tag{2.21}$$

The mean and covariance of the Gaussian approximated filtered distribution are now computed as:

$$\begin{aligned}\mathbf{K}_k &= \mathbf{C}_k \mathbf{S}_k^{-1},\\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}),\\ \mathbf{P}_{k|k} &= \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^{\mathsf{T}}.\end{aligned} \tag{2.22}$$

Due to the employment of sigma-points in the UKF to capture the nonlinear behaviour of the motion and measurement models, the algorithm is technically a sample-based filter. The sample is however not random since the locations of the sigma-points in the state-space are distinctly defined by the UT, cf. equation (2.13). By this strategic selection, only a few sigma-points—and associated weights—are required to effectively compute the first and second moments of the filtering distribution; as opposed to the required size of a random sample.

Since the goal of the UKF algorithm in this thesis is to approximate the filtering distribution as a Gaussian distribution, represented only by the mean and the covariance, some information stored in the sigma-points is lost between iterations. The following section will present filtering methods in which the information contained in the sample is instead retained throughout the complete estimation sequence. This approach results in more versatile and accurate algorithms, but also increases the computational cost.

2.4 Monte Carlo Methods

Monte Carlo methods refer to a class of techniques employed to estimate statistical quantities from random samples of a probability distribution. The convergence of such methods is guaranteed by the central limit theorem [39], i.e., estimates approach their true values as the sample size $N \to \infty$.

In Bayesian filtering, the objective to compute the filtered density, cf. equations (2.5) and (2.6), can often be reduced to evaluating expectations of some transformed random variable, conditioned on observed data. That is, given some arbitrary function, $\mathbf{g} : \mathbb{R}^n \mapsto \mathbb{R}^m$, the problem is to compute integrals of the form:

$$\mathrm{E}\left[\mathbf{g}(\mathbf{x}) \mid \mathbf{y}_{1:T}\right] = \int \mathbf{g}(\mathbf{x}) p(\mathbf{x} \mid \mathbf{y}_{1:T}) \,\mathrm{d}\mathbf{x}, \tag{2.23}$$

where $p(\mathbf{x} \mid \mathbf{y}_{1:T})$ is the posterior probability distribution of the the target quantity $\mathbf{x}$ given all measurements $\mathbf{y}_1, \ldots, \mathbf{y}_T$.

Monte Carlo methods can be employed to numerically approximate integrals of the form (2.23). Given a set of random samples $\{\mathbf{x}^{(i)} : i = 1, \ldots, N\}$ from the probability distribution of interest $p(\mathbf{x} \mid \mathbf{y}_{1:T})$, the best estimator for the expectation is the average:

$$\mathrm{E}\left[\mathbf{g}(\mathbf{x}) \mid \mathbf{y}_{1:T}\right] \approx \frac{1}{N} \sum_{i=1}^{N} \mathbf{g}(\mathbf{x}^{(i)}). \tag{2.24}$$

2.4.1 The Particle Filter

A particle filter employs sequential Monte Carlo methods to approximate the solution to the optimal filtering equations (2.5) and (2.6). The particle idea amounts to propagating through time a set of weighted particles $\{(w_k^{(i)}, \mathbf{x}_k^{(i)}) : i = 1, \ldots, N\}$ drawn from an *importance distribution*, representing the filtering distribution $p(\mathbf{x}_k \mid \mathbf{y}_{1:k})$. This is in contrast to, e.g., the KF and the UKF, which condenses all information about the filtering density into its mean $\hat{\mathbf{x}}_{k|k}$ and covariance $\mathbf{P}_{k|k}$ in each iteration. The particle filter also differs in that the measurement update amounts to updating the particle weights $w^{(i)}$ instead of the actual state realisations $\mathbf{x}^{(i)}$. Particle filters are thus very flexible and therefore particularly suitable for highly nonlinear problems and are capable of characterising multimodal distributions.

A significant problem that arises in the particle filter is that of particle *degeneracy*. Sequential importance sampling is a basic particle filter where, in each filter iteration, samples are drawn from the importance distribution and weights are calculated recursively to approximate the posterior probabilities of the samples. Since the information provided by observations are injected into the estimate through the weights, w_k, it is possible for the algorithm to reach a point where almost all weights are equal or close to zero; under the condition that $\sum_{i=1}^{N} w_k^{(i)} = 1$. This means that, even though using a large number of samples, the particle filter cannot produce a useful estimate. The solution to the degeneracy problem is referred to as *resampling*, and can be summarised as the following algorithm:

Algorithm 3 (Resampling)**.** *Given a set of weighted particles* $\{(w_k^{(i)}, \mathbf{x}_k^{(i)}) : i = 1, \ldots, N\}$ *at time step k, resampling is performed by*

1. *Generating indices* $j \in \{1, \ldots, N\}$ *such that* $p(j = i) = w_k^{(i)}$ *for each* $i = 1, \ldots, N$.
2. *Replacing the old particle set with* $\mathbf{x}_k^{(j)}$*; duplicating probable particles and removing degenerate particles.*
3. *Setting all new weights* $w_k^{(i)} = 1/N$.

Another problem arising in the particle filter scheme is that the filtering distribution is not always such that samples can be drawn from it directly. To overcome this, samples are instead obtained from an importance distribution $\pi(\mathbf{x} \mid \mathbf{y}_{1:T})$ that approximates $p(\mathbf{x} \mid \mathbf{y}_{1:T})$; this is called *importance sampling*. In this thesis, a particular type of particle filter will be investigated, the *bootstrap filter* [44], in which the importance distribution is represented by the dynamical model $p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$ for the process of interest.

Although this choice of $\pi(\mathbf{x} \mid \mathbf{y}_{1:T})$ is intuitively a sensible idea, the dynamical model turns out to be a generally inefficient choice for the importance distribution [39]. As such, the bootstrap filter usually requires a large set of particles to accurately approximate the filtering distribution, as well as resampling of the particles at each iteration.

Algorithm 4 describes the bootstrap filter, using resampling to handle the degeneracy problem.

Algorithm 4 (Bootstrap filter)**.** *The bootstrap filter is initialised by drawing N samples $\mathbf{x}_0^{(i)}$ from the prior distribution*

$$\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0), \quad i = 1, \ldots, N,$$

and setting all weights $w_0^{(i)} = 1/N, i = 1, \ldots, N$. For each iteration $k = 1, \ldots, T$, do the following:

1. *Draw new samples $\mathbf{x}_k^{(i)}$ from the importance distribution, represented by the dynamic model:*
$$\mathbf{x}_k^{(i)} \sim p(\mathbf{x}_k \mid \mathbf{x}_{k-1}^{(i)}), \quad i = 1, \ldots, N.$$
2. *Calculate the weights for each particle from the likelihood model of the measurements:*
$$w_k^{(i)} \propto p(\mathbf{y}_k \mid \mathbf{x}_k^{(i)}), \quad i = 1, \ldots, N,$$
and normalise them to sum to unity.
3. *Perform resampling of the particles, according to Algorithm 3.*

The filtering distribution of particle filters is approximated by a sum of weighted Dirac delta functions, as

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:k}) \approx \sum_{i=1}^{N} w_k^{(i)} \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)}). \tag{2.25}$$

However, since resampling is performed in every iteration of the bootstrap filter, the particle weights are all identical: $w_k^{(i)} = 1/N$ for all $i = 1, \ldots, N$. Thus, the approximated first and seconds moments of the filtered density (2.10) are simply obtained as the sample average and covariance of the particle set; such that (2.10) is described by

$$\hat{\mathbf{x}}_{k|k} \approx \bar{\mathbf{x}}_{k|k} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_k^{(i)}, \tag{2.26a}$$

$$\mathbf{P}_{k|k} \approx \bar{\mathbf{P}}_{k|k} = \frac{1}{N-1} \sum_{i=1}^{N} (\mathbf{x}_k^{(i)} - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k^{(i)} - \hat{\mathbf{x}}_{k|k})^\mathsf{T}. \tag{2.26b}$$

2.4.2 The Ensemble Kalman Filter

The EnKF is a sampling based nonlinear filter with similarities to the linear KF, the UKF, and the PF. The algorithm differs from other nonlinear filters most importantly in that it is derived as an approximation to the linear KF, in contrast to actually attempting to solve the nonlinear filtering problem.

The Ensemble Idea and Algorithm

The EnKF centres around the idea to propagate an *ensemble* of state realisations $\{\mathbf{x}_k^{(i)} : i = 1, \ldots, N\}$; similar to the PF and in contrast to the UKF, where only the first two moments of the filtering distribution are stored between filter iterations [32]. The estimate

and associated uncertainty of the filtered density can thus in any iteration be obtained by computing the average and covariance of the state ensemble, see (2.26).

For the formulation of the EnKF algorithm it is convenient to collect the ensemble of state realisations in a matrix $\mathbf{X}_k \in \mathbb{R}^{n\times N}$, whose columns are the state vectors $\mathbf{x}_k^{(i)}, i = 1, \dots N$. A natural ensemble extension of the process noise $\mathbf{q}_k$ in the nonlinear filtering model (2.11) is to assign a Gaussian noise vector to each state realisation, creating the noise matrix $\mathbf{V}_k$. Using the same notation for sample mean and covariance as in equations (2.26), the state ensemble of deviations from the mean is defined as

$$\widetilde{\mathbf{X}}_k \triangleq \mathbf{X}_k - \bar{\mathbf{x}}_k \mathbb{1}^\mathsf{T} = \mathbf{X}_k(\mathbf{I}_N - \tfrac{1}{N}\mathbb{1}\mathbb{1}^\mathsf{T}),$$

where $\mathbb{1} \triangleq [1, \dots, 1]^\mathsf{T}$ is a vector of ones. Analogous notation is used for an ensemble of measurement predictions $\mathbf{Y}_k$ and its associated noise matrix $\mathbf{E}_k$; refer to the nonlinear problem in (2.11). The EnKF is summarised in Algorithm 5.

Algorithm 5 (Ensemble Kalman filter)**.** *The ensemble Kalman filter for nonlinear problems of the form* (2.11) *is initialised by drawing N samples* $\mathbf{x}_0^{(i)}$ *from the prior distribution*

$$\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0), \quad i = 1, \dots, N,$$

to form the initial state ensemble matrix $\mathbf{X}_{0|0}$*. For each iteration* $k = 1, \dots, T$*, perform the following prediction and update steps:*

Prediction: *Form the process noise matrix* $\mathbf{V}_{k-1}$ *by drawing N samples from the process noise distribution,*

$$\mathbf{V}_{k-1} = \begin{bmatrix} \mathbf{q}_{k-1}^{(1)} & \cdots & \mathbf{q}_{k-1}^{(N)} \end{bmatrix}, \quad \mathbf{q}_{k-1}^{(i)} \sim \mathrm{N}(\mathbf{0}, \mathbf{Q}_{k-1}), \quad i = 1, \dots, N,$$

and propagate the state ensemble through the dynamic model to obtain the ensemble of state predictions:

$$\mathbf{X}_{k|k-1} = \mathbf{f}(\mathbf{X}_{k-1|k-1}) + \mathbf{V}_{k-1}. \tag{2.27}$$

Update: *Form the measurement noise matrix* $\mathbf{E}_k$ *by drawing N samples from the measurement noise distribution,*

$$\mathbf{E}_k = \begin{bmatrix} \mathbf{r}_k^{(1)} & \cdots & \mathbf{r}_k^{(N)} \end{bmatrix}, \quad \mathbf{r}_k^{(i)} \sim \mathrm{N}(\mathbf{0}, \mathbf{R}_k), \quad i = 1, \dots, N,$$

and compute the ensemble of output predictions by propagating the ensemble of state predictions through the measurement model,

$$\mathbf{Y}_{k|k-1} = \mathbf{h}(\mathbf{X}_{k|k-1}) + \mathbf{E}_k. \tag{2.28}$$

Compute the state and measurement ensembles of deviations,

$$\begin{aligned} \widetilde{\mathbf{X}}_{k|k-1} &= \mathbf{X}_{k|k-1}(\mathbf{I}_N - \tfrac{1}{N}\mathbb{1}\mathbb{1}^\mathsf{T}), \\ \widetilde{\mathbf{Y}}_{k|k-1} &= \mathbf{Y}_{k|k-1}(\mathbf{I}_N - \tfrac{1}{N}\mathbb{1}\mathbb{1}^\mathsf{T}), \end{aligned} \tag{2.29}$$

and obtain the updated state ensemble from the sample covariance of the measurements $\bar{\mathbf{S}}_k$ *and the sample cross-covariance of the state and measurement* $\bar{\mathbf{C}}_k$ *as:*

$$\begin{aligned} \bar{\mathbf{S}}_k &= \tfrac{1}{N-1}\widetilde{\mathbf{Y}}_{k|k-1}\widetilde{\mathbf{Y}}_{k|k-1}^{\mathsf{T}}, \\ \bar{\mathbf{C}}_k &= \tfrac{1}{N-1}\widetilde{\mathbf{X}}_{k|k-1}\widetilde{\mathbf{Y}}_{k|k-1}^{\mathsf{T}}, \\ \bar{\mathbf{K}}_k &= \bar{\mathbf{C}}_k\bar{\mathbf{S}}_k^{-1}, \\ \mathbf{X}_{k|k} &= \mathbf{X}_{k|k-1} + \bar{\mathbf{K}}_k(\mathbf{y}_k\mathbb{1}^{\mathsf{T}} - \mathbf{Y}_{k|k-1}). \end{aligned} \tag{2.30}$$

The EnKF is not restricted to Gaussian initial state and noise distributions, $p(\mathbf{x}_0)$, $p(\mathbf{q}_k)$, $p(\mathbf{r}_k)$, neither is the noise required to be additive. This means that Algorithm 5 can be generalised to handle a wider range of filtering problems [32]. However, for the purpose of this thesis, these distributions are all assumed to be Gaussian; as is common in SOC estimation, e.g., [17]–[20], [24], [25], [45], [46].

Important Advantages and Disadvantages

The main advantage of the EnKF, and the motivation for its development, is its ability to handle extremely high-dimensional estimation problems. In cases where n is very large, other algorithms suffer from the need to store an $n \times n$ covariance matrix of the filtered distribution $\mathbf{P}_{k|k}$. In contrast, the EnKF can be implemented with an ensemble size $N < n$ to circumvent this problem. Comparing with the linear KF, the EnKF can in this situation reduce the number of floating point operations required by a factor N/n [32].

Using a too small ensemble size N, in particular when $N < n$, has proven to result in severe underestimations of the state uncertainty $\mathbf{P}_{k|k}$ [47]–[49]. Since the ensemble is initially a random sample of the state distribution, in contrast to the UKF, it is not guaranteed that the N points in the state-space are able to adequately describe the density. This effect persists as the ensemble is transformed by the possibly nonlinear motion and measurement functions, resulting in an over-confident estimate $\mathbf{X}_{k|k}$. Furthermore, as the Kalman gain is computed from the ensemble itself, see (2.30), the ensemble can not be considered a random sample after a few iterations of the algorithm. The update step couples the ensemble members nonlinearly to each other; which results in a sample that is neither independent nor Gaussian [32].

Convergence of the mean and covariance, computed from the ensemble in the EnKF, towards the KF estimate in (2.9), is only guaranteed for linear filtering problems, as $N \to \infty$. For nonlinear problems, the EnKF does not give the Bayesian filtering solution, as concluded in [50].

The discussion so far has concerned problems where n is large, $n > N$. To describe the battery dynamics considered in the SOC estimation problem it is not necessary to use a large number of states, which means that it is affordable to use a large N, $N \gg n$, in the EnKF implementation. This will result in an algorithm that is very similar to the PF, with the important exception that the state realisations are directly updated, without association to particle weights.

3

Lithium-Ion Battery Modelling

This chapter introduces the fundamental concepts of Lithium-ion battery modelling in the framework of the low-complexity equivalent-circuit cell models. The set of electrical analogies presented here, for describing the otherwise electrochemical behaviour of lithium-ion battery cells, are in Chapter 4 used to formulate the estimation problem.

3.1 A Physical Perspective

Energy is stored chemically in a battery cell by supplying an electrical current, due to the transportation of lithium from the cathode (positive electrode) to the anode (negative electrode). When a current is demanded from the battery, lithium is instead moved in the opposite direction, discharging the battery. This is illustrated in Figure 3.1. The concept of the cell's state-of-charge is thus directly related to the concentration of lithium in the solid particles c_s of the cell anode.

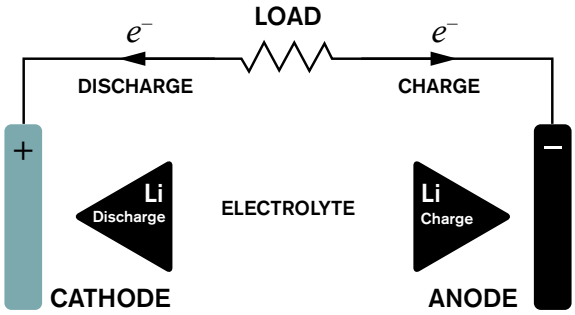


Figure 3.1: Conceptual illustration of the electrochemical mechanics in a lithium-ion battery cell.

The physical definition of state-of-charge is based on the stoichiometry

$$\alpha = \frac{c_{s,\text{avg}}}{c_{s,\text{max}}}, \tag{3.1}$$

which is the ratio of the present average lithium concentration in the anode to the possible maximum equivalent. The cell state-of-charge z at any time t can hereby be defined as

$$z(t) = \frac{\alpha(t) - \alpha_0}{\alpha_{100} - \alpha_0}, \tag{3.2}$$

where α_{100} and α_0 corresponds to the ratio (3.1) when the cell is fully charged and fully discharged, respectively.

3.2 Equivalent-Circuit Cell Models

Modern lithium-ion batteries are complex electrochemical systems that are difficult and cumbersome to adequately describe from a physical perspective. For the problem addressed in this thesis, it turns out to be sufficient—and necessary for implementation in real-time applications—to consider the input–output behaviour of a single battery cell as a *black-box* system. In the context of black-box modelling, the input–output system behaviour is analysed to infer the parameters of some ready-made model; refer to Section 4.2.2 for a more detailed description of these models. Generally, these parameters do not have any physical interpretation.

Electrical components are commonly used to represent the observed static and dynamic electrical behaviour, resulting in an Equivalent Circuit Model (ECM). For instance, an electrical resistor represents an internal voltage drop and power loss within the battery cell, due to the passage of current, while a capacitor models an intrinsic voltage inertia. In this case, the system input is the load current,

$$i(t) \begin{cases} > 0 & \text{during discharge,} \\ < 0 & \text{during charge,} \end{cases} \tag{3.3}$$

and the system output is the cell terminal voltage, $v(t)$.

Results from investigating the single cell can be generalised to different types of battery packs, as suggested in [5], and will not be treated here.

3.2.1 Open-Circuit Voltage

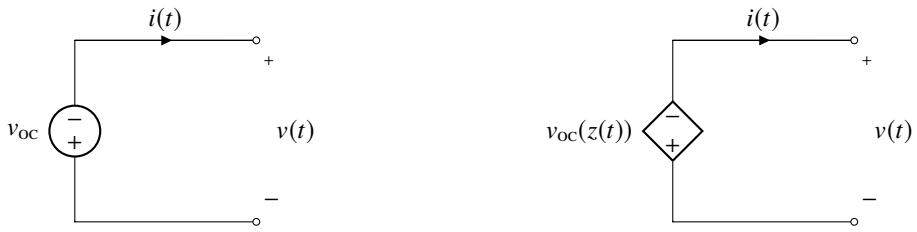
The Open-Circuit Voltage (OCV) refers to the effective voltage across the terminal of the circuit under a no-load condition. An ideal battery can be considered as an ideal voltage source, as demonstrated in Figure 3.2a. However, in reality, a battery cell's OCV tends to decrease as it is being discharged, i.e., it is dependent on the residual capacity or equivalently the SOC; this is illustrated in Figure 3.2b. Furthermore, the OCV also depends on its operating temperature.

The SOC and temperature dependencies are both nonlinear, but static, i.e. there is a direct correspondence[1] between a certain value of SOC and OCV, and likewise for temperature. These relationships are commonly determined from careful cell tests in a lab environment, as in [1]. The data can then be used to create a multidimensional Look-Up Table (LUT) or to fit polynomials of suitable order; approximating the relationships. The latter method is often used in connection with the EKF as filtering algorithm, due to the need to compute derivatives for linearisation of the motion and measurement models.

3.2.2 Battery Cell Capacity

In Section 3.1 the SOC was defined in a physical context. The remainder of this chapter will almost exclusively consider the battery cell from a strictly electrical perspective. The battery is primarily a vessel for storing electrical energy. To continue the discussion in the

[1] There is a hysteresis present in the OCV as a function of SOC for charging and discharging of the battery cell.



(a) OCV represented by an ideal voltage source. **(b)** OCV as a function of SOC, $z(t)$.

Figure 3.2: Two equivalent circuits modelling the OCV of a battery cell.

following section, the *total capacity* of the cell is here formally defined, as well as when a cell is considered to be fully charged and fully discharged. Last, we make an electrical definition of SOC.

Definition 1 (Total capacity)**.** The quantity of charge removed from a cell as it is brought from a fully charged state to a fully discharged state is termed the *total capacity*, Q.

Definition 2 (Fully charged cell)**.** A cell is fully charged when its OCV reaches $v_h(T)$, a manufacturer specified voltage that may be a function of temperature T.

Definition 3 (Fully discharged cell)**.** A cell is fully discharged when its OCV reaches $v_l(T)$, a manufacturer specified voltage that may be a function of temperature T.

Definition 4 (Residual capacity)**.** The quantity of charge that would be removed from a cell if it were brought from its present state to a fully discharged state is called the *residual capacity*, Q_{res}.

Definition 5 (State-of-charge)**.** The ratio of the residual capacity to the total capacity of the cell is defined as the cell *state-of-charge*; $z \triangleq Q_{\text{res}}/Q$.

Definitions 1 through 5 are illustrated in Fig. 3.3.

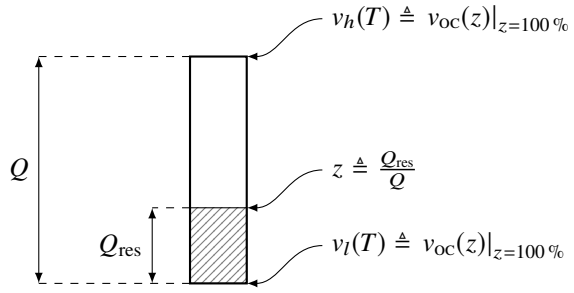


Figure 3.3: Schematic illustration of a single battery cell with definitions of cell total capacity Q, residual capacity Q_{res}, fully charged voltage v_h and fully discharged voltage v_l as functions of temperature T, and SOC, z.

3.2.3 State-of-Charge

From Definition 5 it is clear that SOC changes only due to changes of the residual capacity,[2] which is the accumulated charge in the battery cell due to passage of current,

$$\dot{Q}_{\text{res}}(t) = i(t). \tag{3.4}$$

From Definition 5, using the sign convention for the input current in (3.4), we arrive at the dynamic equation for the SOC in a battery cell as

$$\dot{z}(t) = -\eta \frac{1}{Q} i(t), \tag{3.5}$$

where η is the Coulombic efficiency parameter [1], representing loss of energy when charging the cell.

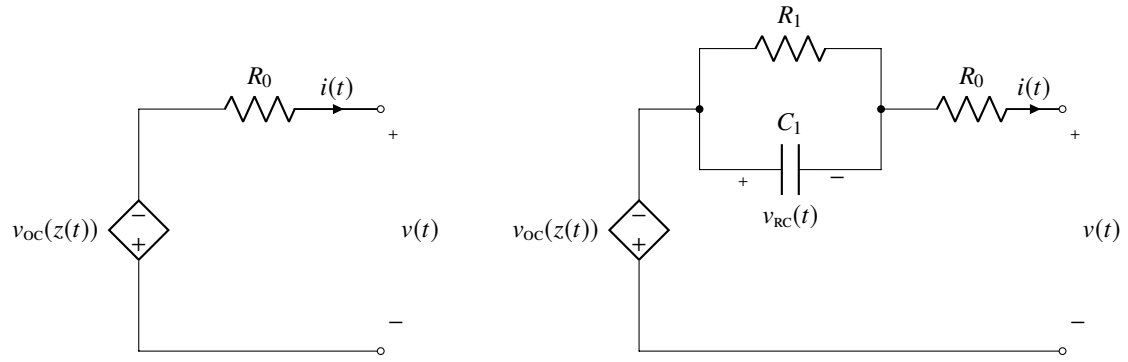
3.2.4 Linear Polarisation and Diffusion Voltages

Electrochemical polarisation in battery cells refers to the internal side-effects in a battery cell from changes in input current; isolating the electrodes from the electrolyte. Polarisation is important in the mathematical abstraction of a battery cell because it results in a departure of the terminal voltage from the open-circuit voltage of the cell, $v < v_{\text{OC}}$. The observed voltage difference,

$$v_d \triangleq v_{\text{OC}}(z) - v, \tag{3.6}$$

is referred to as *diffusion voltage*.

Diffusion voltage is the collective effect of both an instantaneous response to the change in input current as well as a dynamic (non-instantaneous) one.



(a) Instantaneous polarisation modelled by a single resistor in series with the cell OCV.

(b) Extending (a) with an RC sub-circuit, modelling dynamic polarisation.

Figure 3.4: Two equivalent circuits modelling the input–output behaviour of battery cell polarisation.

[2]This is true if the cell total capacity is regarded as a constant, time-invariant quantity. In reality, the total capacity will slowly change as the battery is ageing, but this effect is negligible in the context of this thesis.

Equivalent Series Resistance

Voltage drop across an equivalent series resistance, R_0, effectively models the instantaneous diffusion voltage. In an actual circuit, power would dissipate through the internal resistance as heat, and as such the internal resistance models the static inefficiency of the battery cell at load. The instantaneous diffusion model is illustrated by the circuit in Figure 3.4a, where the input–output behaviour can be obtained using Kirchhoff's Voltage Law (KVL),

$$v(t) = v_{\mathrm{OC}}(z(t)) - i(t)R_0. \tag{3.7}$$

Dynamic Polarisation

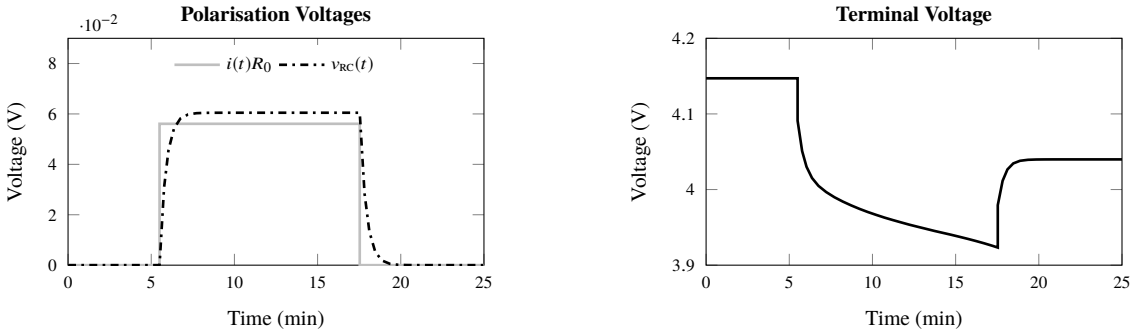
When the cell is allowed to rest, terminal voltage of the cell gradually approaches OCV. This phenomenon is due to the slow diffusion processes occurring in the cell, observable as a slowly changing voltage. The dynamic behaviour, referred to as dynamic diffusion voltage, can be closely approximated by one or more parallel RC sub-circuits. In Figure 3.4b, the static polarisation model described in (3.7) has been extended by one such RC branch. Using KVL, we obtain the dynamic input–output voltage behaviour

$$v(t) = v_{\mathrm{OC}}(z(t)) - v_{\mathrm{RC}}(t) - i(t)R_0, \tag{3.8}$$

where

$$\dot{v}_{\mathrm{RC}}(t) + \frac{1}{R_1 C_1} v_{\mathrm{RC}}(t) = \frac{1}{C_1} i(t). \tag{3.9}$$

Figure 3.5 displays example simulation results from the first order RC model in Figure 3.4b. To illustrate the contribution from the static and dynamic polarisation models respectively, a discharge pulse of 1C-magnitude[3] is applied to the circuit between $t = 5.5$ min and $t = 17.5$ min. The polarisation voltage drop is plotted in Figure 3.5a and the output terminal voltage in Figure 3.5b.



(a) Cell polarisation per model component.

(b) Cell output voltage.

Figure 3.5: Simulation results of a first order RC model subjected to a pulse current between time instances $t = 5.5$ min and $t = 17.5$ min. Note the drop in steady-state terminal voltage, due to the dependence of OCV on SOC.

[3] A measure of battery cell discharge rate, expressed relative to the maximum cell capacity. The 1C-rate for a cell with a total capacity of 20 Ah is 20 A, which will discharge the cell completely in one hour.

4

Implementation and Methods

This chapter presents the mathematical equations for the motion and measurement models of the nonlinear SOC filtering problem. Identification of equivalent circuit parameters for the first order RC model are also described for two different lithium-ion battery cell chemistries; denoted E2 and A123. Furthermore, the data sets and test scenario are explained, as well as a systematic method to measure filter performance.

4.1 The Estimation Problem

Model-based estimation problems require well-defined motion and measurement models. Estimation is carried out for two cases: state-only estimation and state-and-parameter estimation. In the first case, the motion model for the battery dynamics can be described with a linear transition function in time. In the second case, however, when the model parameters are assumed to be non-stationary, the joint motion model for the evolution of states and parameters is nonlinear. The explicit mathematical expressions for these two different cases are given below.

We define the state vector, $\mathbf{x}_k \in \mathbb{R}^n$, to contain the SOC and the voltage drop across the RC sub-circuit in the ECM,

$$\mathbf{x}_k \triangleq \begin{bmatrix} z_k & v_k^{\mathrm{RC}} \end{bmatrix}^\mathsf{T}, \tag{4.1}$$

and the vector of model parameters to be estimated, $\boldsymbol{\theta}$, to contain the equivalent series resistance, R_0, and the dynamic polarisation parameters, R_1 and C_1, as described in equations (3.8) and (3.9) and Figure 3.4b,

$$\boldsymbol{\theta}_k \triangleq \begin{bmatrix} R_{0,k} & R_{1,k} & C_{1,k} \end{bmatrix}^\mathsf{T}. \tag{4.2}$$

The measurement, $\mathbf{y}_k \in \mathbb{R}^m$, providing observations of the real process, is defined to be the battery cell terminal voltage,

$$\mathbf{y}_k \triangleq v_k. \tag{4.3}$$

The states, parameters, and measurements are all associated with uncertainties, modelled using additive white noise processes, explained in chapter 2, and are denoted $\mathbf{q}_k$, $\mathbf{q}_k^{\boldsymbol{\theta}}$, and $\mathbf{r}_k$ respectively. Furthermore, we define a deterministic input, $\mathbf{u}_k \in \mathbb{R}^p$, to the system of interest as the load current, representing the energy demand of an EV,

$$\mathbf{u}_k \triangleq i_k. \tag{4.4}$$

As the input is considered to be deterministic, no uncertainty is associated with the battery cell load current, i_k.

4.1.1 State-Only Estimation

When assuming stationary battery model parameters, the target quantity of interest in the estimation scheme is the state vector $\mathbf{x}$, defined in (4.1): the cell SOC and the voltage drop across the RC subcircuit. As such, the motion model is linear in its states, while the measurement model depends nonlinearly on the SOC. We can describe this generically as

$$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{q}_{k-1}, \tag{4.5a}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{r}_k, \tag{4.5b}$$

where $\mathbf{A}_{k-1} \in \mathbb{R}^{n\times n}$ and $\mathbf{B}_{k-1} \in \mathbb{R}^{n\times p}$ are the parameter dependent state transition and control matrices, and $\mathbf{h} : \mathbb{R}^n \mapsto \mathbb{R}^m$ is the nonlinear measurement model. The two matrices, $\mathbf{A}_k$ and $\mathbf{B}_k$, can be obtained by discretization of the continuous-time differential equations (3.5) and (3.9). Assuming a sampling time T_s, equation (4.5a) can be written explicitly as

$$\begin{bmatrix} z_k \\ v_k^{\mathrm{RC}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & e^{\left(-\frac{T_s}{R_1 C_1}\right)} \end{bmatrix} \begin{bmatrix} z_{k-1} \\ v_{k-1}^{\mathrm{RC}} \end{bmatrix} + \begin{bmatrix} -T_s\eta/Q \\ R_1\left(1 - e^{\left(-\frac{T_s}{R_1 C_1}\right)}\right) \end{bmatrix} i_{k-1} + \mathbf{q}_{k-1}. \tag{4.6}$$

The state vector (4.1) is mapped to the measurements (4.3) according to the first order RC model voltage equation (3.8). The measurement model (4.5b) is thus explicitly defined as

$$v_k = v_{\mathrm{oc}}(z_k) - v_k^{\mathrm{RC}} - R_0 i_k + \mathbf{r}_k. \tag{4.7}$$

4.1.2 State-and-Parameter Estimation

Owing to the electrochemical dynamics of the battery cells, the battery model parameters $\boldsymbol{\theta}$, defined in (4.2), varies over a charging or discharging cycle. To account for this, parameters are jointly estimated by augmentation of the state vector, enabling a possibly more accurate estimation of the state variable of interest, z. A simple way to model the parameter evolution, when there is no meaningfully identifiable behaviour, is by assuming a Gaussian random walk, i.e., the parameters are assumed to be constant, but driven by a noise process $\mathbf{q}_k^{\boldsymbol{\theta}}$.

The generic motion model for the state-and-parameter estimation case can be described as

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \boldsymbol{\theta}_{k-1}, \mathbf{u}_{k-1}) + \mathbf{q}_{k-1}, \tag{4.8a}$$

$$\boldsymbol{\theta}_k = \boldsymbol{\theta}_{k-1} + \mathbf{q}_{k-1}^{\boldsymbol{\theta}}, \tag{4.8b}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta}_k, \mathbf{u}_k) + \mathbf{r}_k, \tag{4.8c}$$

where the state transition function $\mathbf{f} : \mathbb{R}^n \mapsto \mathbb{R}^n$ is linear in its states, but nonlinear when considering an augmented state-and-parameter vector $[\mathbf{x}, \boldsymbol{\theta}]^\mathsf{T}$. The explicit expression for (4.8a) is the same as (4.6), but with time-varying model parameters, $\boldsymbol{\theta}$. Likewise, the measurement model (4.8c) in the state-and-parameter estimation scheme is the same as (4.5b), but is now nonlinearly dependent on the model parameters.

4.2 Parametric System Identification

The first order RC model, illustrated in Figure 3.4b and characterised by equations (3.8) and (3.9), is an equivalent circuit model and does not, as such, have any physical meaning in terms of its electrical parameters,

$$\boldsymbol{\theta} = \begin{bmatrix} R_0 & R_1 & C_1 \end{bmatrix}^{\mathsf{T}}.$$

In order to choose a suitable black-box model structure to ultimately identify $\boldsymbol{\theta}$, a brief system analysis is here presented for describing the input–output behaviour of the circuit, see Figure 3.4b, in terms of its parameters. This is followed by a description and selection of model structure as well as the method used to fit model parameters to measurement data. The section is concluded with a comparison between the data and the resulting model predictions.

4.2.1 System Analysis

The continuous-time transfer function from load current $i(t)$ to total diffusion voltage,

$$v_d(t) = v_{\text{RC}}(t) + i(t)R_0,$$

where $v_{\text{RC}}(t)$ is defined in eq. (3.9), is obtained by Laplace transformation:

$$\frac{V_d(s)}{I(s)} = R_0 + \frac{R_1}{1 + R_1C_1s} = \frac{R_0 + R_1 + R_0R_1C_1s}{1 + R_1C_1s}. \tag{4.9}$$

Here, s denotes the complex Laplace transform variable. For computer implementation a discrete-time transfer function is required. Here, one is obtained using the bilinear transform,

$$s = \frac{2}{T_s}\frac{1 - z^{-1}}{1 + z^{-1}}, \tag{4.10}$$

where z now denotes the discrete-time equivalent of the variable s, and T_s is the sampling time. Using (4.10), the discrete-time transfer function that corresponds to (4.9) is

$$\frac{V_d(z)}{I(z)} = \frac{\beta_0 + \beta_1 z^{-1}}{1 + \alpha_1 z^{-1}}, \tag{4.11}$$

with coefficients defined in terms of the equivalent electrical parameters as

$$\alpha_1 = \frac{T_s - 2R_1C_1}{T_s + 2R_1C_1}, \tag{4.12a}$$

$$\beta_0 = \frac{R_0T_s + R_1T_s + 2R_0R_1C_1}{T_s + 2R_1C_1}, \tag{4.12b}$$

$$\beta_1 = \frac{R_0T_s + R_1T_s - 2R_0R_1C_1}{T_s + 2R_1C_1}. \tag{4.12c}$$

Given knowledge about the coefficients α_1, β_0, and β_1, the RC model parameters $\boldsymbol{\theta}$ can be obtained from the expressions in (4.12) as

$$R_0 = \frac{\beta_0 - \beta_1}{1 - \alpha_1}, \tag{4.13a}$$

$$R_1 = \frac{2\beta_1 - 2\beta_0\alpha_1}{1 - \alpha_1^2}, \tag{4.13b}$$

$$C_1 = \frac{(1 - \alpha_1)^2}{4(\beta_1 - \beta_0\alpha_1)}. \tag{4.13c}$$

4.2.2 Linear Time-Series Models

The objective here is to find the, in some sense, best possible values for the coefficients a_1, b_0, and b_1 for the discrete-time transfer function in (4.11). These will statically describe the battery system in state-only estimation, and serve as the initial guess in the state-and-parameter estimation scheme. For the particular problem of finding black-box parameter values, a common method is to consider observed input, $u(t)$, and output, $y(t)$, sequences for the system in eq. (4.11). Since actual measurements are utilised, it is also natural to assume that a, possibly coloured, error signal, $e(t)$, is present in the data. This scenario can be written generally in the time-domain using the shift operator q [51], which is such that $qy(t_k) = y(t_{k+1})$ where $t_k = kT_s$, as

$$y(t) = G(q, \boldsymbol{\vartheta})u(t) + H(q, \boldsymbol{\vartheta})e(t), \tag{4.14}$$

Here, the input is the load current, $u(t) \equiv i(t)$, the output is the diffusion voltage, $y(t) \equiv v_d(t)$, $G(q, \boldsymbol{\vartheta})$ represents the transfer function in (4.11), $H(q, \boldsymbol{\vartheta})$ is the error signal linear filter, and $\boldsymbol{\vartheta}$ denotes the linear time-series model parameters, as coefficients of the transfer functions $G(q)$ and $H(q)$. Obtaining the best parameter values depends on how the structure of $G(q)$ and $H(q)$ are selected.

The transfer functions in the general form (4.14), called the Box-Jenkins (BJ) model, can be written as polynomial rational functions of the shift-operator q:

$$G(q, \boldsymbol{\vartheta}) = \frac{B(q)}{F(q)} = \frac{b_0 + b_1 q^{-1} + b_2 q^{-2} + \cdots + b_{n_b} q^{-n_b}}{1 + f_1 q^{-1} + f_2 q^{-2} + \cdots + f_{n_f} q^{-n_f}}, \tag{4.15a}$$

$$H(q, \boldsymbol{\vartheta}) = \frac{C(q)}{D(q)} = \frac{c_0 + c_1 q^{-1} + c_2 q^{-2} + \cdots + c_{n_c} q^{-n_c}}{1 + d_1 q^{-1} + d_2 q^{-2} + \cdots + d_{n_d} q^{-n_d}}, \tag{4.15b}$$

where $F(q)$ is the auto-regressive part of the output dynamics, $B(q)$ describes the injection of the exogenous input into the system, $C(q)$ represents a moving average function of the white noise signal, and $D(q)$ represents the dynamics that the disturbance signal is subjected to when entering the system. A summary of the three most relevant linear time-model structures are presented in Table 4.1: the Auto Regressive Moving Average with eXogenous input (ARMAX) model, the Auto Regressive with eXogenous input (ARX) model, and the Output-Error (OE) model. The main problem to consider when selecting the polynomials in (4.15) is how to model the injection of the disturbance signal $e(t)$ into the system.

For the problem at hand, it is reasonable to assume that the error enters early in the process [2] and that it is thus coloured by the auto-regressive part of the output dynamics.

Table 4.1: Summary of relevant linear time-series model structures. The polynomial time-shift transfer functions are divided into the following categories: the auto regression of the output $A(q)y(t)$, the exogenous input $B(q)u(t)$, and the moving average of the white noise $C(q)e(t)$ [51]; cf. equations (4.14) and (4.15).

Model Type	Mathematical Structure	Description
ARMAX	$F(q) = D(q) = A(q)$ $A(q)y(t) = B(q)u(t) + C(q)e(t)$	*Auto-Regressive Moving Average with eXogenous input*: Noise and input are subjected to the same dynamics. Suitable if noise enters early in the process.
ARX	$C(q) \equiv 1$ $A(q)y(t) = B(q)u(t) + e(t)$	*Auto Regressive with eXogenous input*: Simpler form of the ARMAX model, without the moving average function of the white noise.
OE	$H(q) \equiv 1$ $A(q)y(t) = B(q)u(t) + A(q)e(t)$	*Output Error*: The properties of the disturbance signals are not modelled, and the error is simply the difference between the actual output and the noise-free output.

The simplest model structure, among those in Table 4.1, satisfying this assumption is the ARX model [51]. The mathematical significance of this choice, referring to equations (4.14) and (4.15), is that $F(q) = D(q) = A(q)$ and $C(q) \equiv 1$, meaning that the noise is strongly influenced by the auto-regressive operator of the battery diffusion voltage dynamics [2]. This is illustrated in Figure 4.1.

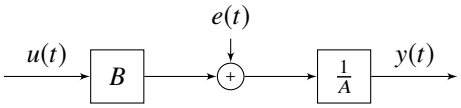


Figure 4.1: Block diagram describing the ARX model structure.

Thus, we arrive at the following linear time-series model for the current to diffusion voltage dynamics:

$$A(q)y(t) = B(q)u(t) + e(t). \tag{4.16}$$

With model orders $n_a = 1$ and $n_b = 2$, as suggested by the discrete-time transfer function in equation (4.11), we have that (4.16) is equivalent to

$$y(t_k) + a_1 y(t_{k-1}) = b_0 u(t_k) + b_1 u(t_{k-1}) + e(t_k). \tag{4.17}$$

4.2.3 Linear Regression

To find the best vector of parameters $\boldsymbol{\vartheta}^*$, the ARX linear-in-parameters system output *predictor* $\hat{y}(t \mid \boldsymbol{\vartheta})$ is constructed from (4.17), using discrete measurement sequences of the input $u(t)$ and output $y(t)$ collected at uniformly spaced time instants t_k, $k = 1, \ldots, N$, as

$$\hat{y}(t_k \mid \boldsymbol{\vartheta}) = -a_1 y(t_{k-1}) + b_0 u(t_k) + b_1 u(t_{k-1}) = \boldsymbol{\vartheta}^\mathsf{T} \boldsymbol{\varphi}(t_k), \tag{4.18}$$

where

$$\boldsymbol{\vartheta} = \begin{bmatrix} a_1 & b_0 & b_1 \end{bmatrix}^\mathsf{T}, \quad \text{and} \quad \boldsymbol{\varphi} = \begin{bmatrix} -y(t_{k-1}) & u(t_k) & u(t_{k-1}) \end{bmatrix}^\mathsf{T}. \tag{4.19}$$

The ARX model prediction error, given the set of identified parameters $\boldsymbol{\vartheta}$, is then

$$\varepsilon(t_k \mid \boldsymbol{\vartheta}) = y(t_k) - \hat{y}(t_k \mid \boldsymbol{\vartheta}). \tag{4.20}$$

By forming the mean squared error

$$V_N(\boldsymbol{\vartheta}) = \frac{1}{N} \sum_{k=1}^{N} \varepsilon^2(t_k \mid \boldsymbol{\vartheta}), \tag{4.21}$$

an estimate of the best vector of parameters is obtained, in the least-squares sense, by solving the minimisation problem

$$\boldsymbol{\vartheta}_N^* = \arg\min_{\boldsymbol{\vartheta}} V_N(\boldsymbol{\vartheta}). \tag{4.22}$$

Using MATLAB's System Identification Toolbox, an ARX model of order $n_a = 1$ and $n_b = 2$ was fitted with a congruence of 97 % to current and voltage measurement data for the E2 and A123 battery chemistries. Model predictions are compared to measurements in Figure 4.2. Observation of the diffusion voltage, used as output in the system identification process, were derived from the relationship (3.6) using LUTs for mapping SOC to OCV and a sequence of true SOC values. See Section 4.3 for a detailed explanation of this data.

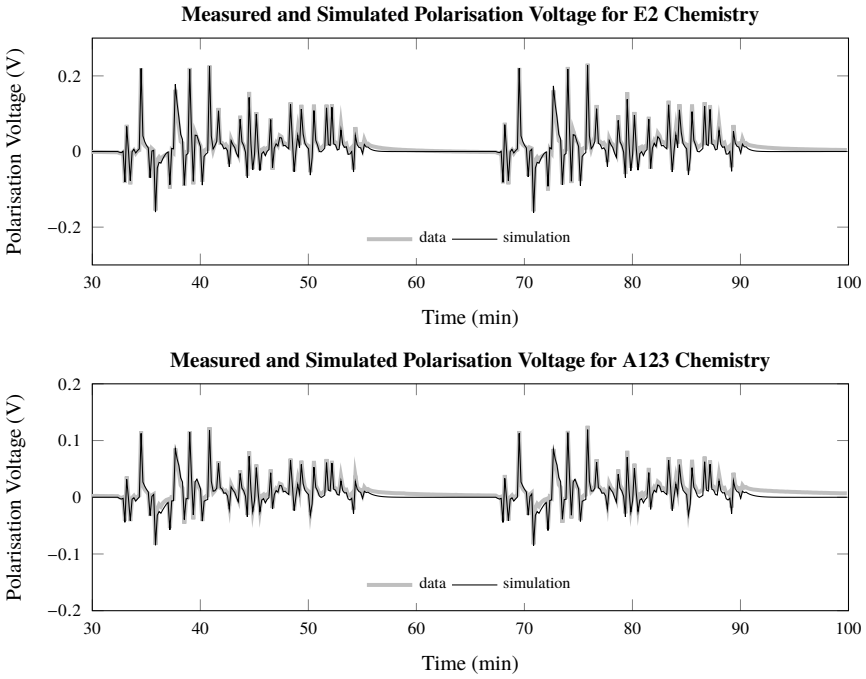


Figure 4.2: System identification results for two different lithium-ion cell chemistries: E2 (top) and A123 (bottom). The first 30–100 min are shown for the measured data, $v_d = v - v_{\text{oc}}$, 25 °C, and an ARX(1,2,0) model output prediction. The first 30 min not include in the plot are used to bring the battery cell to its nominal operating mode. A fitness of 97 % is achieved for both chemistries.

The values for the best vector of parameters are given in Table 4.2, expressed in terms of $\boldsymbol{\theta}$.

4.2.4 Significance of Non-Stationary Parameters

One of the key problems in SOC estimation is that the systems are non-stationary. This means that no single set of parameters $\boldsymbol{\vartheta}$ can describe the complete battery cell input–output

Table 4.2: Results from system identification for the first order RC model.

Chemistry	R_0 [mΩ]	R_1 [mΩ]	C_1 [F]
E2	10.8	12.3	1941
A123	10.8	11.6	1979

behaviour as the SOC changes. The proposed solution to this problem investigated in this thesis is to jointly estimate how the model parameters change when SOC changes. For this purpose, the system identification is made using only an initial sub-sequence of the input–output data, where the process is nearly stationary. The identified parameters will then serve as an initial guess for the state-and-parameters estimation scheme. The same initial-guess parameter values are used for the state-only estimation algorithm as well, to illustrate the significance of the adaptive approach.

4.3 Data Sets and Test Scenario

Data sets for algorithm benchmarking and validation are obtained from the supplementary material to [1], available[1] at [52], and consists of:

- Time signals for cell load current, terminal voltage, and SOC when subjected to a typical automotive application load profile at different temperatures, see Figures 4.3 and 4.4.
- Temperature dependent parameter values for coulombic efficiency η and total cell capacity Q, see Figure 4.5.
- OCV curves as functions of SOC, Figure 4.6.

4.3.1 Dynamic Test Sequence

The test scenario is based on an Urban Dynamometer Drive Schedule (UDDS) profile, designed to excite battery cell dynamics in a way representative for a typical EV/HEV application [19]. In the test the battery cell, with initial SOC $z_0 = 100\,\%$, is subjected to a load current consisting of a sequence of UDDS profiles, see Figure 4.3a, alternated with 5 min intervals of zero load current. The UDDS cycle is only applied if the SOC is in the range $z \in [90, 10]\,\%$, which is the estimation region of interest.[2] The entire dynamic test is explained by Program 1.

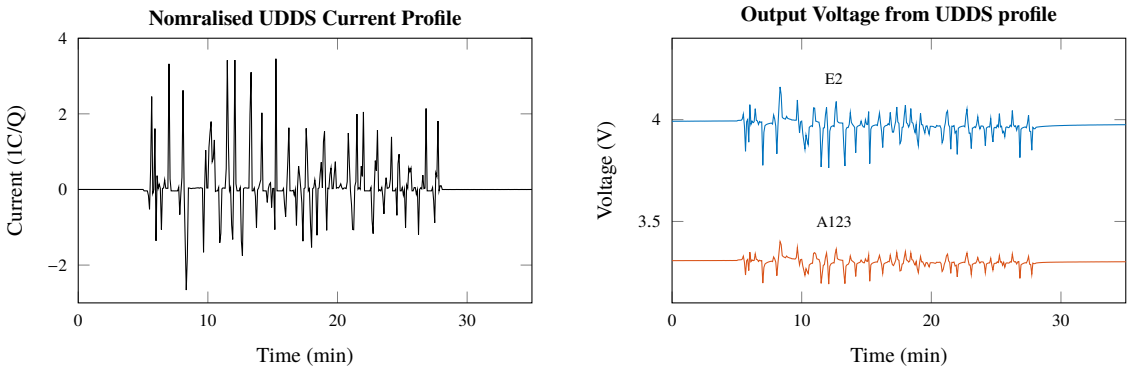
The ground truth for estimation tests is constituted by input–output behaviour obtained from the UDDS tests, illustrated in Figure 4.3, as well as a predetermined "true" SOC signal, see Figure 4.4. As seen in Figure 4.4, the test sequence begins at 100 % SOC. However, to properly evaluate the convergence of the filter SOC estimate to its true value, a filtering sequence is initialised with an offset associated with a relatively high uncertainty. This is further detailed in Section 4.4 and Table 4.3.

[1] Copyright © 2015 by Gregory L. Plett of the UCCS. Licensed under CC BY-NC-SA 4.0.

[2] The SOC range $z \in [90, 10]\,\%$ represents the expected design limits for the cell chemistries of interest [18].

Program 1 Dynamic UDDS discharge test.

Require: Uniform cell test temperature T
Require: $v = v_h(T)$ at steady state: $z = 100\,\%$
Ensure: $z < 10\,\%$
  while $z > 90\,\%$ **do**
    $i \leftarrow$ 1C rate
  end while
  while $z \geq 10\,\%$ **do**
    $i \leftarrow$ UDDS cycle
    if UDDS ended **then**
      $i \leftarrow 0$
      Pause(5 min)
    end if
  end while



(a) One cycle of the UDDS current profile. The desired 1C rate is adjusted for a specific cell by multiplication with its total capacity Q.

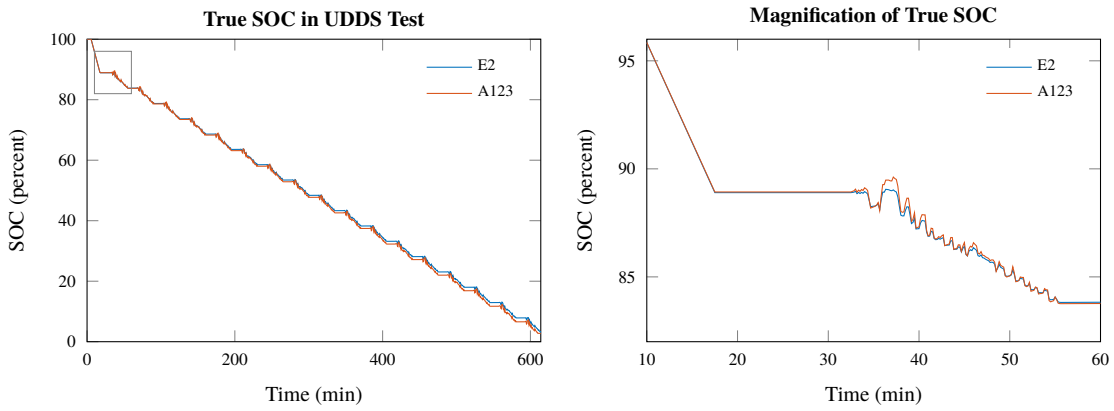
(b) Sample terminal voltage outputs for the E2 and A123 cell chemistries when subjected to the UDDS profile.

Figure 4.3: Sample input and output sequences for the exactly 35 min long UDDS profile.

4.3.2 Temperature Dependent Parameters and Hysteresis

The temperature dependency of total cell capacity Q and coulombic efficiency η, introduced in Chapter 3, are displayed in Figure 4.5 for two different battery cell chemistries: E2 and A123. Parameter values are given for the eight uniformly spaced temperatures $T \in \{-15, -5, 5, 15, 25, 35, 45\}$ °C. However, the available data sets do not provide a meaningful way of varying cell temperature for any single test sequence.

Temperature dependent OCV–SOC curves are also accessible for the same range of temperatures as η and Q. However, due to the complex behaviour of the OCV–SOC hysteresis nonlinearity, and the simplistic first order RC model used in the filtering algorithms, a LUT for *average* OCV is utilised, based on the present SOC estimate and temperature. Furthermore, the nominal desired operating temperature for battery packs in EV/HEV applications is in the range 15–35 °C [53], [54], in which the OCV–SOC relationship is near stationary. Therefore, tests will be restricted to the cell temperature 25 °C, as marked in Figure 4.5, and the OCV–SOC data is incorporated in the estimation scheme as the average of the charge and discharge curves, displayed in Figure 4.6.



(a) True SOC sequence.

(b) Magnification.

Figure 4.4: Ground truth of SOC from the UDDS test in Program 1. To the left (a) is the complete time plot and to the right (b) is a magnified segment between 10 min and 60 min; marked in (a) with a rectangle.

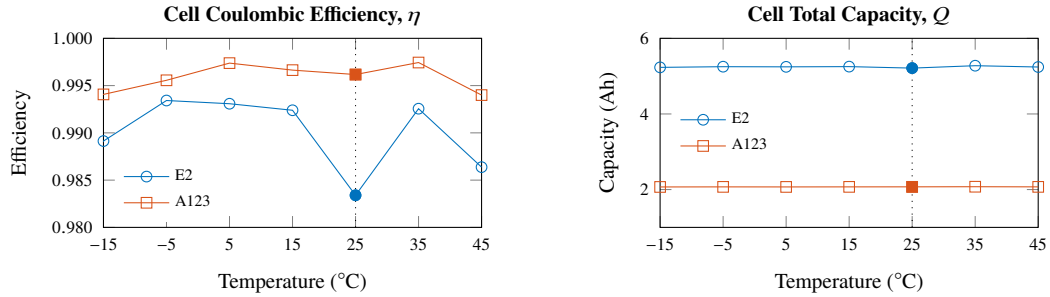


Figure 4.5: Temperature dependent battery cell parameters. Pre-estimated cell coulombic efficiency η and total cell capacity Q for eight different temperatures. The filled markers, at 25 °C, are the parameter values used in the test scenario.

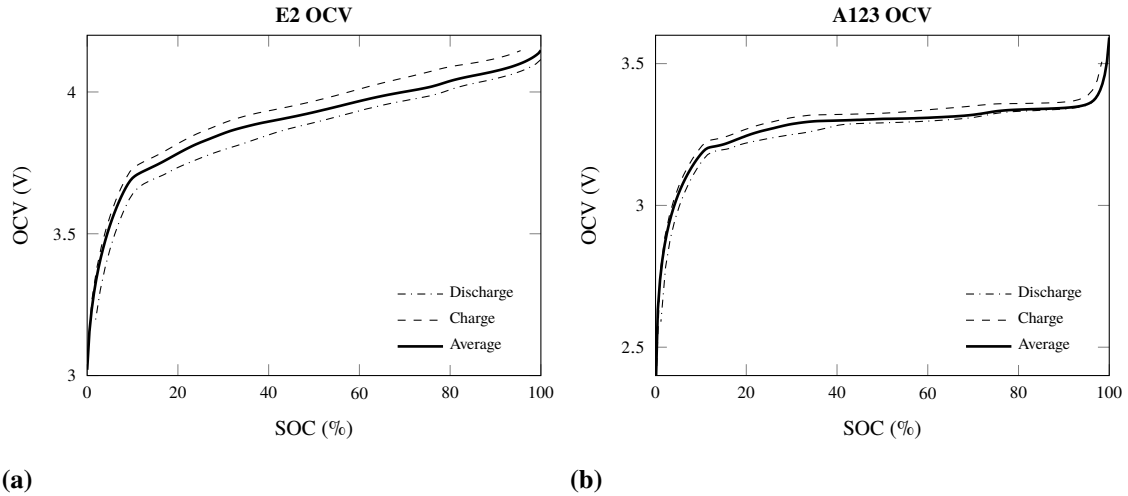
4.4 Filter Dimensions and Tuning

The performance of any filter is highly dependent on proper tuning of its parameters; in particular the noise covariance matrices of the motion and measurement model, denoted $\mathbf{Q}$ and $\mathbf{R}$ respectively. Each application requires a different setting, which depends on the type of processes involved in the estimation problem, the sampling time used in computer implementations, and the level of certainty associated with the mathematical model of the process and the measurements. A brief description of the different tunable parameters involved in this thesis is given below in terms of what they represent in the context of mathematical modelling.

The values used in subsequent tests for the tuning parameters explained below are summarised in Table 4.3.

Initial State Vector $\mathbf{x}_0$

The initial state vector, $\mathbf{x}_0$, represents the prior information about the initial condition of the system. In the dynamic test considered, the battery is being discharged from its fully charged state from rest. This means the initial SOC is 100 % and no voltage drop across the RC sub-circuit. Initial values for parameters R_0, R_1, and C_1 are identified using parametric



(a) **(b)**

Figure 4.6: OCV curves as functions of SOC for 25 °C. The hysteresis observed in charge and dischare for both chemistries is approximated by an average LUT.

system identification explained in Section 4.2. With the purpose to investigate the transient behaviour of the filter, the SOC is initialised with an offset of 20 %. This makes it possible to study the stability and convergence of the filter as the dynamic test progresses.

Initial Covariance $\mathbf{P}_0$

The initial filter covariance matrix, $\mathbf{P}_0$, represents the uncertainty of the initial state of the system in terms of statistical variance. For a system in which the initial state is rather unknown, corresponding values in the $\mathbf{P}_0$ matrix are set to a high value, and conversely to a low value if the initial state is certain. If there's no meaningful way of determining the initial state, $\mathbf{x}_0$, the initial state covariance is usually tuned to represent a higher uncertainty than the measurements, to achieve faster initial convergence of the estimate towards what can be inferred from the available measurements. In the EnKF, the $\mathbf{P}_0$ matrix determines the spread of the initial state realisations, $\mathbf{X}_0$.

Process Noise Q

The process noise matrix, $\mathbf{Q}$, models the uncertainty of the states in the motion model. The diagonal elements represent the intensity, in terms of variance, of the random noise processes that are part of the state evolution. Off-diagonal elements represents any possible correlation between the stochastic parts of state evolutions. For a system with independent states, the off-diagonal elements are usually set to zero. Smaller variance levels in the diagonal of $\mathbf{Q}$ means less uncertainty of a particular state variable, i.e, the state evolution is more deterministic.

For the problem at hand, the evolution of the dynamical system is rather small during each sampling period, $T_s = 1$ s. As such, the stochastic part of the dynamical model is of a smaller or equal magnitude. The order of magnitude for the process noise is determined by analysis of the observable states in the data sets. Further tuning is performed with the objective to improve the SOC estimate.

Measurement Noise R

The measurement noise matrix, **R**, models the measurement uncertainty in terms of the expected variance of the noise in sensor measurements. A lower measurement variance represents higher accuracy of the sensor measurements.

The Kalman gain is computed based on the process noise **Q** and the measurement noise **R**, representing the quality of the information provided by the dynamical model and the observations of the system, respectively. Lower values in **Q** means that more trust is put in the motion model, while lower values in **R** results in a filter that trusts the measurements more. The relative values of **Q** and **R** determine the Kalman gain, **K**, and directly affect the measurement update.

EnKF Ensemble Size N

The number of ensemble members for each state realisation is called the ensemble size. As suggested in Section 2.4.2, there is a lower limit for the ensemble size below which the filter is unable to converge. The number of members in the ensemble also determines the accuracy of the estimate and the associated uncertainty; at an expense of computation time. For lower ensemble sizes, an increase in N would drastically improve the accuracy. However, for larger ensemble sizes, a similar increase in N would not significantly improve the accuracy, but would increase the computation time, leading to a saturation in the effect of the ensemble size. Thus, as long as convergence of the estimate is achieved, the ensemble size can be tuned for the trade-off between accuracy and computational cost.

PF Particle Set Size N

Similar to the EnKF, the PF uses a weighted set of N samples, or particles, to represent the filtering distribution. Analogous arguments can be made regarding the particle size as with ensemble size in respect to convergence, accuracy and computation time. However, since the bootstrap version of the PF is used, the number of particles is required to be higher, relative to the number of ensemble members.

UKF Parameters

As described earlier, the UKF is a sigma-point method, in which the set of sigma-points are chosen deterministically to capture the nonlinearities in the motion and measurement models. The parameters α and κ together define the spread of the sigma-points, while β is related to the prior distribution. Setting the β parameter to 2 is optimal for a prior that is a Gaussian distribution, which is the case here.

4.5 Performance Metrics

A standardised evaluation system is required to assess and gauge the performances of different state estimating algorithms. Different performance metrics—estimation accuracy K_{est}, drift K_{drift}, residual charge determination K_{res}, and transient behaviour K_{trans}, defined in [45] and used in [46]—are incorporated to evaluate each algorithm individually, and to compare the results with other algorithms. Various error boundaries ε are used to define

Table 4.3: Filter tuning parameters for the two different chemistries, E2 and A123. The values for the model parameters R_0, R_1, and C_1 in the initial state vector $\mathbf{x}_0$ have been obtained from the parametric system identification, explained in Section 4.2.

Parameter		E2	A123
Initial state vector, $\mathbf{x}_0$	z	80 %	80 %
	v_{RC}	0 V	0 V
	R_0	10.8 mΩ	10.8 mΩ
	R_1	12.3 mΩ	11.6 mΩ
	C_1	1941 F	1979 F
Initial covariance, $\mathbf{P}_0$	z	5×10^{-3}	5×10^{-3}
	v_{RC}	1×10^{-12}	1×10^{-12}
	R_0	1×10^{-5}	1×10^{-5}
	R_1	1×10^{-5}	1×10^{-5}
	C_1	1×10^{-5}	1×10^{-5}
Process noise, $\mathbf{Q}$	z	1×10^{-9}	1×10^{-9}
	v_{RC}	1×10^{-9}	1×10^{-9}
	R_0	1×10^{-9}	1×10^{-9}
	R_1	1×10^{-9}	1×10^{-9}
	C_1	1×10^{-5}	1×10^{-5}
Measurement noise, $\mathbf{R}$	v	5×10^{-2}	5×10^{-2}
EnKF ensemble size	N	2×10^3	2×10^3
PF particle set size	N	5×10^3	5×10^3
UKF parameters	α	1	1
	β	2	2
	κ	5	5
Sampling time	T_s	1 s	1 s

the performance metrics. Each error boundary corresponds to an evaluation score $P(\varepsilon)$, defined as

$$P(\varepsilon) = \begin{cases} 5, & \text{for} \quad 0\,\% \leq |\varepsilon| \leq 0.5\,\%, \\ 4, & \text{for} \quad 0.5\,\% < |\varepsilon| \leq 1\,\%, \\ 3, & \text{for} \quad 1\,\% < |\varepsilon| \leq 2\,\%, \\ 2, & \text{for} \quad 2\,\% < |\varepsilon| \leq 4\,\%, \\ 1, & \text{for} \quad 4\,\% < |\varepsilon| \leq 8\,\%, \\ 0, & \text{for} \quad |\varepsilon| > 8\,\%. \end{cases} \tag{4.23}$$

4.5.1 Estimation Accuracy K_{est}

Estimation accuracy is the weighted average of evaluation scores for all the error boundaries. Let δ be the absolute difference between the true SOC, z, and the estimated SOC, $\hat{z}$, let $\Sigma\Delta t_{\delta\in\varepsilon_i}$ be total time duration in a certain error boundary, and let t_{total} be the total measured

time for all error boundaries. Then, the estimation accuracy K_{est} is calculated as

$$\begin{aligned} \varepsilon &= z - \hat{z}, \\ K_{\text{est}} &= \sum_{i=1}^{6} \left(P(\varepsilon_i) \frac{\Sigma \Delta t_{\delta \in \varepsilon_i}}{t_{\text{total}}} \right). \end{aligned} \tag{4.24}$$

4.5.2 Drift K_{drift}

Long validation tests or real-time applications during repeated discharge of batteries can lead to error accumulation and drift in the error measurements. Drift in the error can be measured by fitting a regression line,

$$y = at + n, \tag{4.25}$$

to the estimation error, $\varepsilon = z - \hat{z}$. The average drift score, K_{drift}, is calculated by multiplying the gradient a of the regression line with the time representing 1 h of testing, $t_{1\text{h}}$,

$$\varepsilon_{1\text{h}} = a t_{1\text{h}}, \quad K_{\text{drift}} = P(\varepsilon_{1\text{h}}), \tag{4.26}$$

where $\varepsilon_{1\text{h}}$ is the error per hour. This validation principle is illustrated in Figure 4.7.

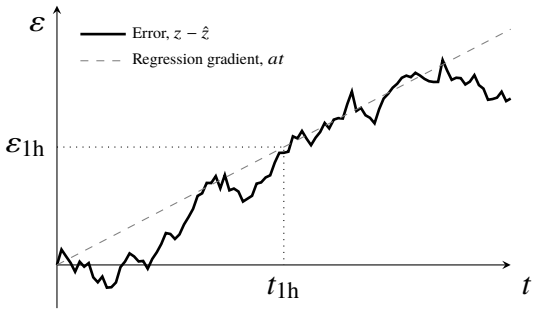


Figure 4.7: Validation principle for the drift error, K_{drift}. The linear regression of the estimation error, $z - \hat{z}$, gives the drift error, $P(\varepsilon_{1\text{h}})$.

4.5.3 Residual Charge Determination K_{res}

The estimation and drift validation metrics are both measures of the overall filtering performance. In general, it is of interest to have access to accurate knowledge about when the battery is starting to deplete. Thus, the estimation scheme is also validated based on the error between the true and the estimated residual capacity, z_{res}. This performance metric is calculated as

$$\begin{aligned} \varepsilon &= |\hat{z}_{\text{res}} - z_{\text{res}}|, \\ K_{\text{res}} &= P(\varepsilon). \end{aligned} \tag{4.27}$$

4.5.4 Transient Behaviour K_{trans}

Transient behaviour is the initial response of the system to any input before it reaches the steady state. Incorrect initial values are the perturbation given to the system to analyse its

transient behaviour. To measure K_{trans}, initial conditions are set to incorrect values and the error bound at the time instant $t_{10\%}$, 10 percent of the total time duration is compared with the difference in true SOC, z, and estimated SOC, $\hat{z}$, at $t_{0\%}$ to evaluate K_{trans} as:

$$\begin{aligned} \varepsilon &= z_{t_{10\%}} - \hat{z}_{t_{10\%}}, \\ K_{\text{trans}} &= P(\varepsilon)\frac{|z_{t_{0\%}} - \hat{z}_{t_{0\%}}|}{z_{t_{0\%}}}. \end{aligned} \tag{4.28}$$

The highest K_{trans} score is achieved when the initial true SOC, z, is 100 % and initial estimated SOC, $\hat{z}$, is 0 %; and the error, ε at $t_{10\%}$ reaches the smallest error bound. This implies that the estimate has managed to reach the true SOC at $t_{10\%}$ from maximum possible initial perturbation.

5
Results

Presented below are results from simulations in MATLAB using recorded measurement data, detailed in Section 4.3, representing the dynamic discharge test sequence outlined in Section 4.3.1. All tests are made assuming a constant temperature of 25 °C, as explained in Section 4.3.2.

Tables 5.1 and 5.2 displays overall performance metrics, defined in Section 4.5, for the three filters in both state-only and state-and-parameter estimation configuration, for tests performed in the two battery cell chemistry types E2 and A123 respectively. The tables also include a set of measures for the SOC estimation error of the entire test sequence. These measures are the Root-Mean-Square Error (RMSE),

$$\mathrm{RMSE}\,(z) = \sqrt{\frac{1}{T}\sum_{k=1}^{T}(z_k - \hat{z}_k)^2}, \tag{5.1}$$

the Mean Absolute Error (MAE),

$$\mathrm{MAE}\,(z) = \frac{1}{T}\sum_{k=1}^{T}|z_k - \hat{z}_k|, \tag{5.2}$$

and the mean filter estimated uncertainty in the form of STandard Deviation (STD),

$$\mathrm{STD}\,(z) = \frac{1}{T}\sum_{k=1}^{T}\sqrt{P^z_{k|k}}, \tag{5.3}$$

where $P^z_{k|k}$ is the estimated variance for the filtered SOC at time step k. The RMSE and MAE have also been computed for the filter output predictions, i.e. the error of the predicted terminal voltage as compared to the corresponding measurements.

Plots displaying evolution of the estimation error and associated estimated uncertainty are provided in Figures 5.1 and 5.2. Parameter evolutions are plotted in Figures 5.4 and 5.5 for the joint filters, as well as in Figure 5.6 in the form of estimated total internal battery cell resistance.

A comparison of the state-only EnKF performance in the A123 chemistry is also made in Figure 5.3 for different ensemble sizes. The performance is evaluated in terms of the estimated uncertainty, related to the actual estimation error, see (5.3). The portion of the total test sequence where the SOC estimate has a larger error than three standard deviations, see (5.3), the estimation accuracy, defined in (4.24); and the total computation time.

Table 5.1: Performance benchmarks for estimation tests performed with the E2 battery cell chemistry type. The SOC estimation errors are given in terms of RMSE, MAE, and the filter estimated STD of the SOC estimate for all three filter algorithms in both state-only and state-and-parameter estimation. Estimation scores are given in terms of the performance metric defined in section 4.5. The error of the filter output predictions are also given, in terms of both RMSE and MAE.

Metric	Unit	UKF	Joint UKF	EnKF	Joint EnKF	PF	Joint PF
SOC RMSE	%	2.13	2.43	1.07	0.95	1.72	2.29
SOC MAE	%	1.93	2.22	0.94	0.79	1.63	2.24
SOC STD	%	0.73	0.86	0.48	0.49	0.32	0.31
Output RMSE	mV	11.20	11.15	13.90	13.86	13.99	16.76
Output MAE	mV	8.29	8.28	10.75	10.68	10.24	11.60
K_{est}	—	2.75	2.57	3.76	3.90	2.88	2.13
K_{drift}	—	5	5	5	5	5	5
K_{res}	—	3	3	3	3	3	3
K_{trans}	—	0.6	0.6	0.8	1	0.6	0.4

Table 5.2: Performance benchmarks for estimation tests performed with the A123 battery cell chemistry type. The SOC estimation errors are given in terms of RMSE, MAE, and the filter estimated STD of the SOC estimate for all three filter algorithms in both state-only and state-and-parameter estimation. Estimation scores are given in terms of the performance metric defined in section 4.5. The error of the filter output predictions are also given, in terms of both RMSE and MAE.

Metric	Unit	UKF	Joint UKF	EnKF	Joint EnKF	PF	Joint PF
SOC RMSE	%	1.43	1.28	0.60	0.57	0.44	0.41
SOC MAE	%	0.83	0.74	0.30	0.28	0.20	0.18
SOC STD	%	0.42	0.45	0.31	0.32	0.27	0.29
Output RMSE	mV	18.22	16.67	31.90	30.97	45.26	43.23
Output MAE	mV	14.22	12.85	23.49	22.72	27.13	27.13
K_{est}	—	4.14	4.19	4.86	4.87	4.93	4.90
K_{drift}	—	5	5	5	5	5	5
K_{res}	—	2	2	2	2	3	3
K_{trans}	—	1	1	1	1	1	1

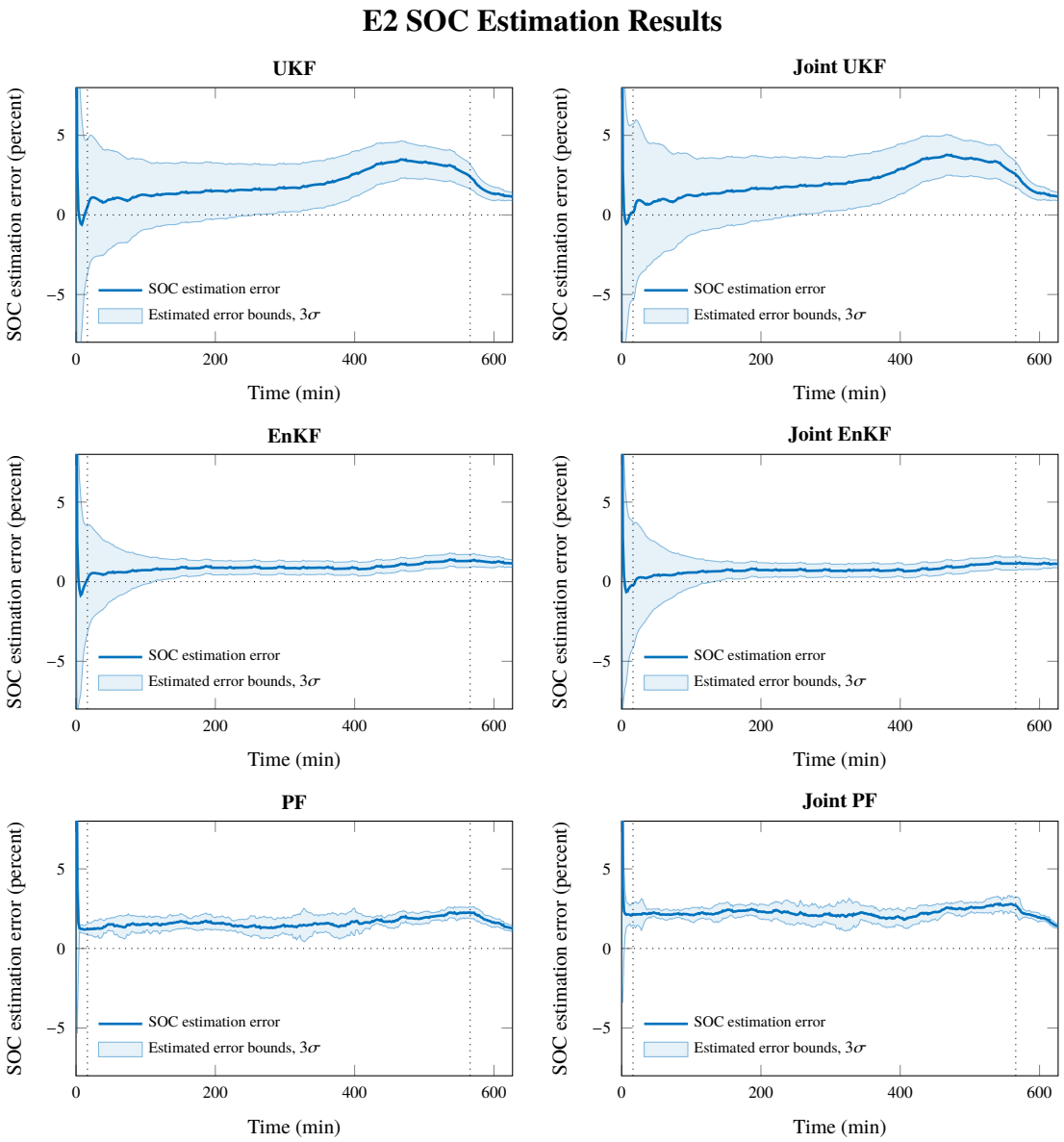


Figure 5.1: Estimation results for the three different filter algorithms, in both state-only and state-and-parameter (joint) estimation configurations, for the UDDS discharge test of the E2 chemistry at 25 °C. The estimation error (thick line) $\Delta z = \hat{z} - z$ is plotted with the associated filter uncertainty (thin lines, shaded area); displayed as three standard deviations from the estimate. The vertical dotted lines mark where the true SOC reaches 90 % and 10 %.

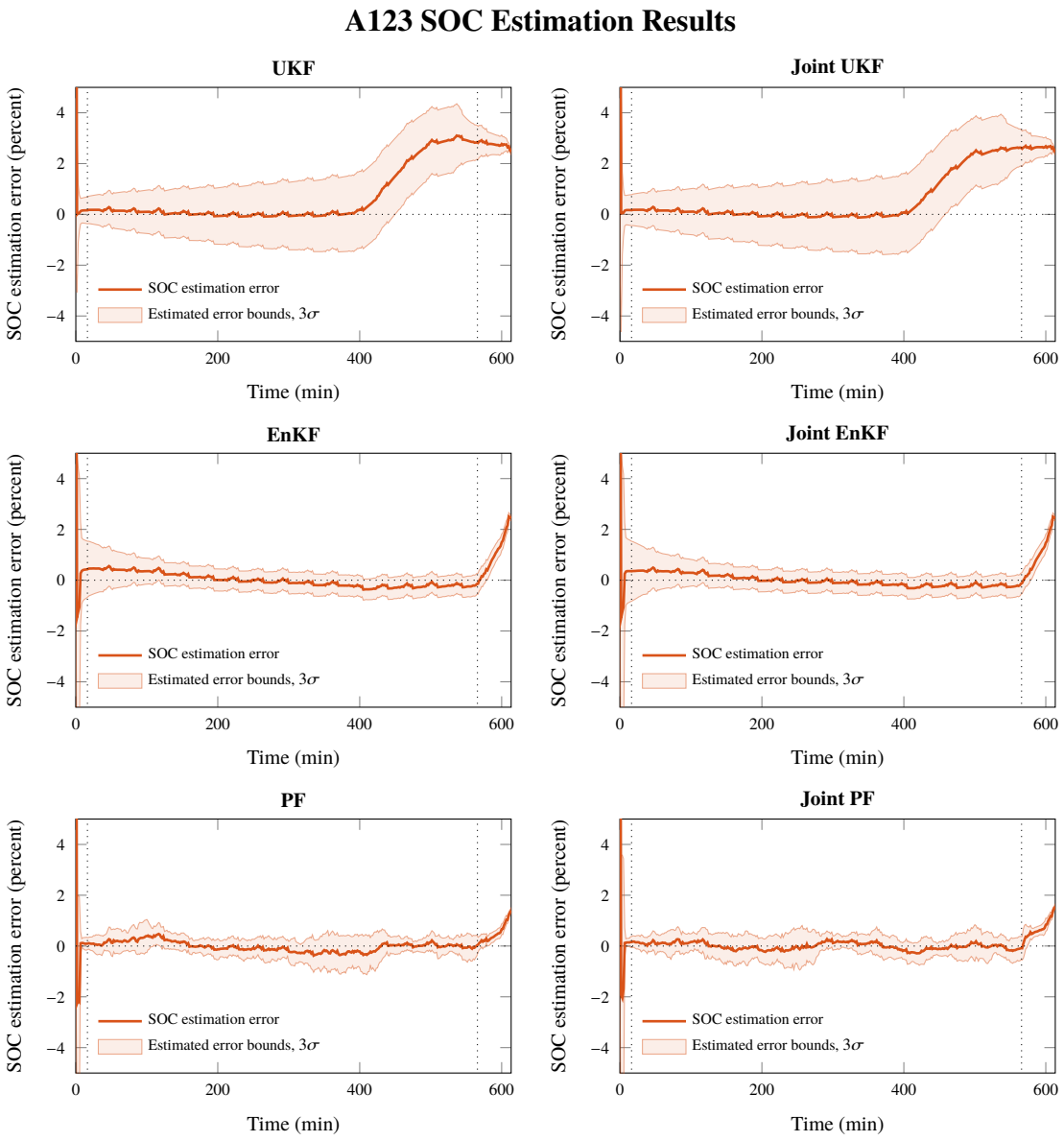


Figure 5.2: Estimation results for the three different filter algorithms, in both state-only and state-and-parameter (joint) estimation configurations, for the UDDS discharge test of the A123 chemistry at 25 °C. The estimation error (thick line) $\Delta z = \hat{z} - z$ is plotted with the associated filter uncertainty (thin lines, shaded area); displayed as three standard deviations from the estimate. The vertical dotted lines mark where the true SOC reaches 90 % and 10 %.

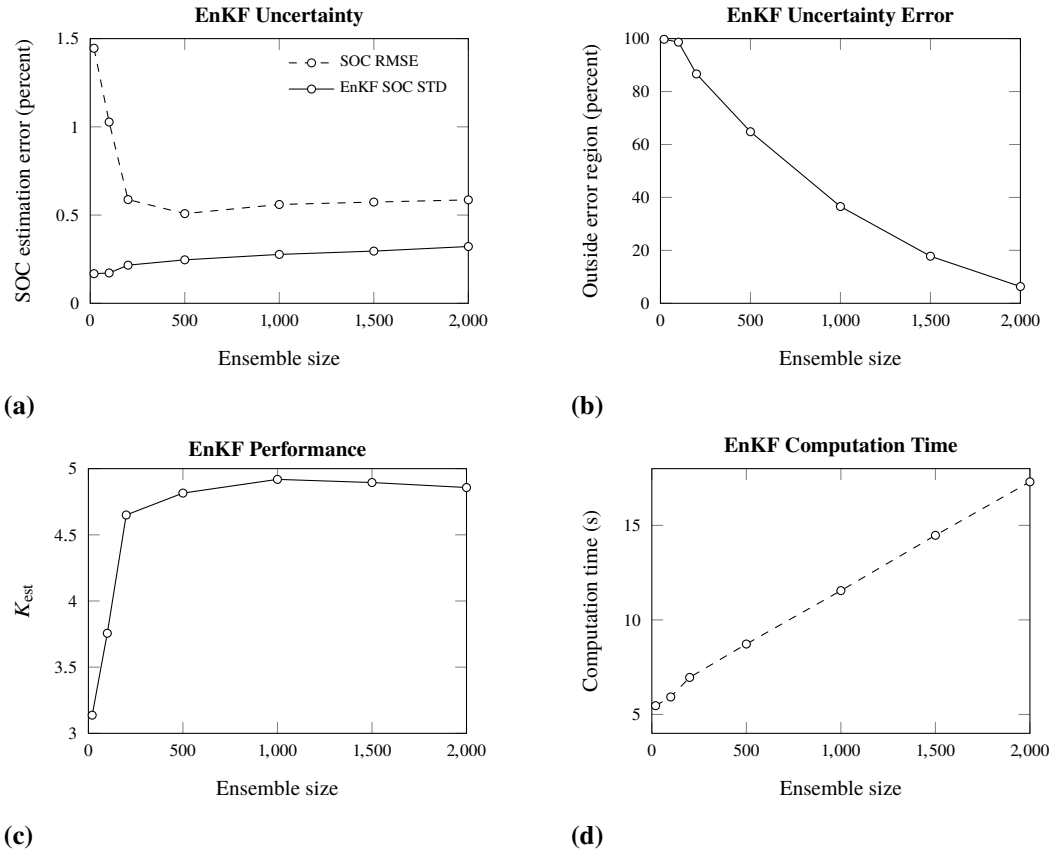


Figure 5.3: Performance metrics for the EnKF using different ensemble sizes for state-only estimation in A123 cell type. In (a) is the SOC RMSE plotted together with the estimated uncertainty as standard deviation. In (b) is the time, in percentage of the total estimation time, plotted for which the absolute estimation error is larger 3 standard deviations; computed from the filter estimated uncertainty. In (c) is the estimation performance metric K_{est} plotted, and in (d) the computation time for the total estimation sequence.

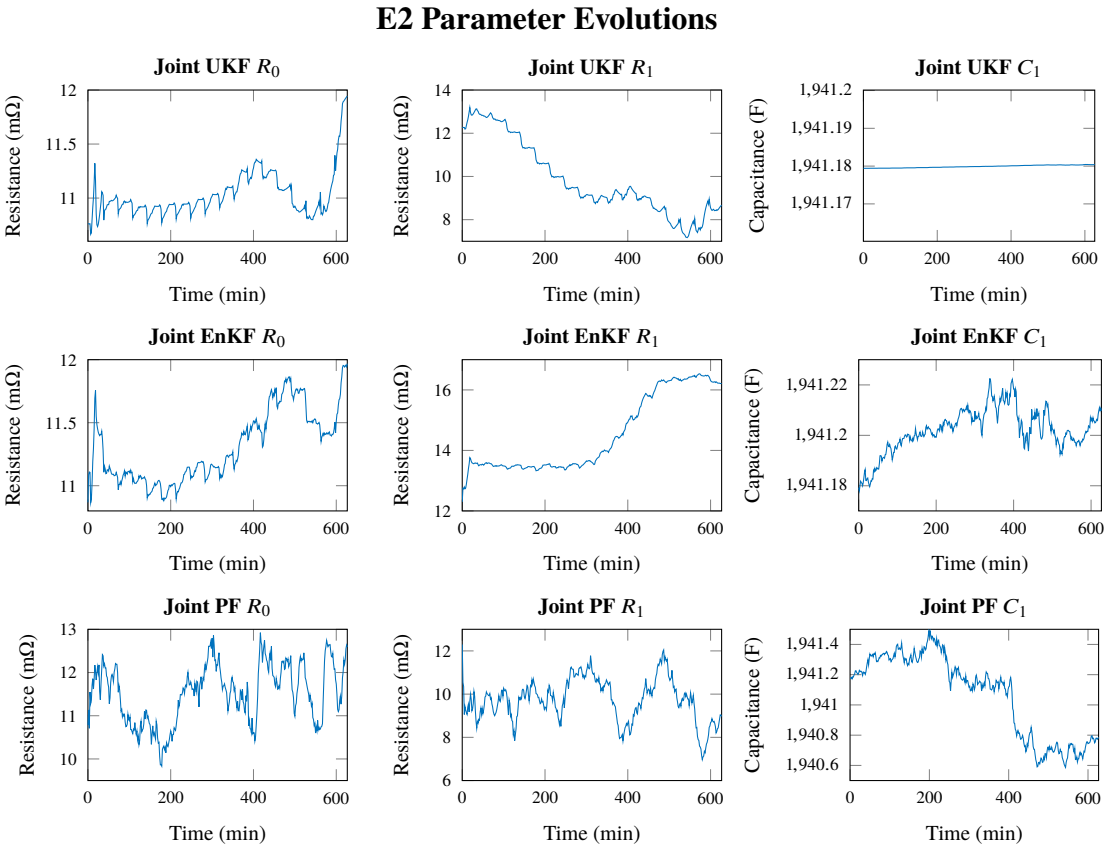


Figure 5.4: Evolutions of estimated ECM parameters for the UDDS discharge test of the E2 chemistry at 25 °C. Note that due to the employment of an ECM, there is no ground truth available for these results.

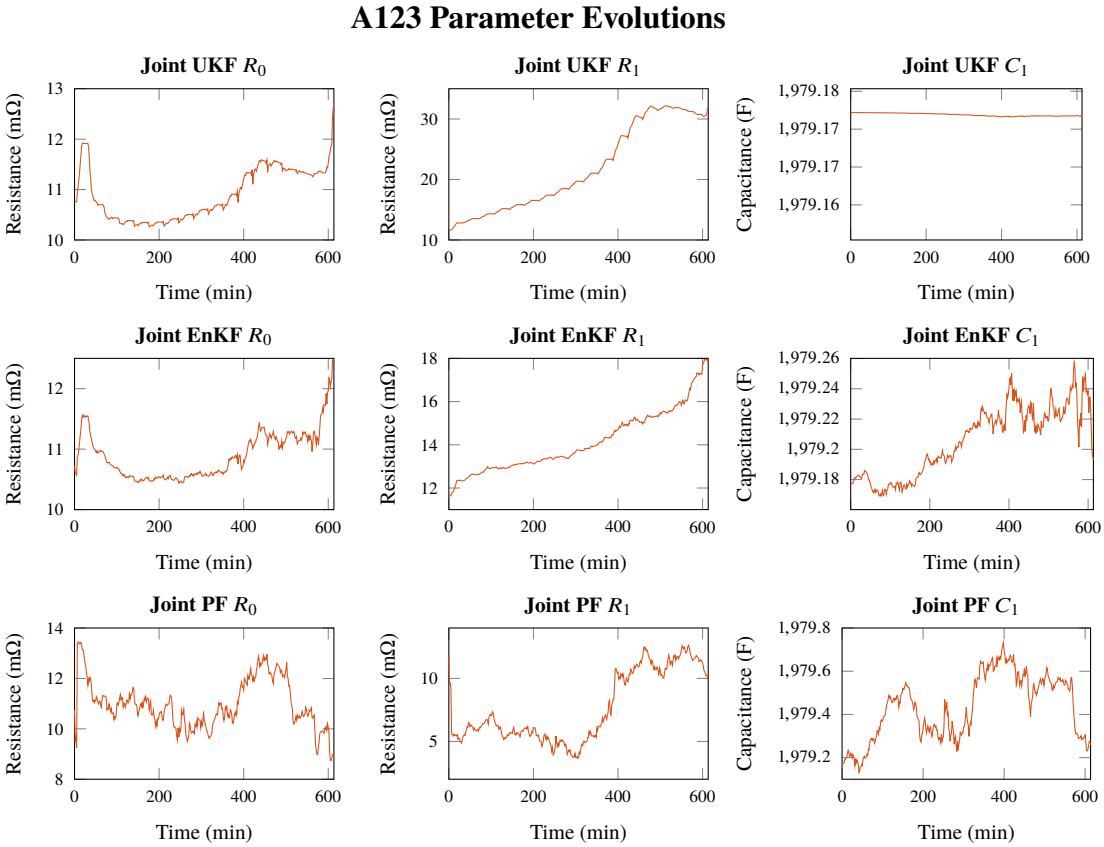


Figure 5.5: Evolutions of estimated ECM parameters for the UDDS discharge test of the A123 chemistry at 25 °C. Note that due to the employment of an ECM, there is no ground truth available for these results.

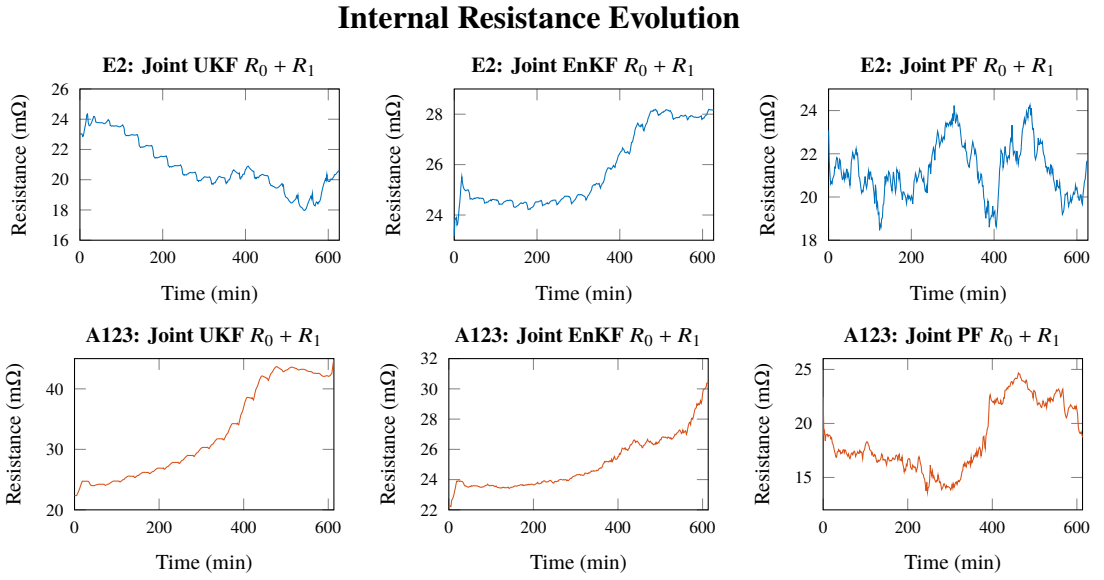


Figure 5.6: Estimated total internal resistance. The sum of the resistance parameters, R_0 and R_1, as an approximation of the total internal resistance of the E2 (top) and A123 (bottom) battery cells.

6
Discussion

The discussion for this thesis begins with a review of the results presented in Chapter 5. Estimation errors for the SOC are discussed and compared for the two different battery cell chemistry types, E2 and A123, as well as in regard to the parameter estimation results. A comparison of the estimation results are also made in relation to an external study.

A discussion is also provided regarding the EnKF dimensions and tuning, followed by evaluations of some of the most important choices made concerning the methodology and scientific approach of this thesis.

6.1 Discussion of Estimation Results

Results using the UKF, EnKF, and PF to estimate the SOC with and without parameters—employing the first order RC model for output predictions—are shown in Figures 5.1 and 5.2; with associated numerical performance metrics in Tables 5.1 and 5.2 respectively. Rather than showing the actual state evolution together with the different estimate sequences, we have plotted the SOC estimation error, $z - \hat{z}$, for the complete dynamic discharge test. We have also plotted the associated uncertainty of the estimate as error bounds corresponding to three STDs, computed from $\mathbf{P}_{k|k}$. Preferably, the zero-error should be included in this confidence region for all estimates, or at least when the true SOC is in the region of interest, which corresponds to 10–90 % SOC. The STD is furthermore an indicator of whether a particular algorithm is working properly or not. Preferably, the average estimated filter STD should be in the same order of magnitude as the estimation RMSE; both metrics are presented in Tables 5.1 and 5.2.

E2 Cell Tests

We begin by discussing the tests performed on the E2 battery cell chemistry, with results displayed in Figure 5.1. The filters are purposefully initialised with an incorrect SOC value of 80 %, the true initial value being 100 %, to study the convergence of the filter estimates to the true value; called transient behaviour. The initial uncertainty is therefore relatively large, but reduces over time and converges to different domains for the different filters. The UKF and EnKF are both able to converge faster to the true state. The PF, however, does not converge, but produces a SOC estimate with a nearly constant offset of about 2 % from the true value throughout the discharge sequence. The convergence of the PF to the wrong value could be the result of erroneous initial parameters, errors in the OCV–SOC LUT, using too few particles, or improper tuning. As the discharge test progresses, the estimation error is kept at a constant level, except for the UKF, which slowly diverges from the zero-error. At the very end of the discharge test the estimation errors of all filters are

reduced to about 1 %, with a comparatively small uncertainty as the predicted terminal voltage approaches v_l, resulting in an intermediate residual score, K_{res}, of 3 for all filters.

The constant offset observed for all filters is likely related to the OCV–SOC relationship, which in reality includes a quite substantial hysteresis for the E2 chemistry, but is approximated by an average curve, see Figure 4.6a. Since both charging and discharging of the battery cell is occurring during the dynamic discharge test, i.e., the sign of the load current is alternating, the average approximation introduces an uncertainty in the measurement function that can yield very different SOC correspondences. For instance, at 60 % SOC, the OCV could vary in 80 mV between the charge and discharge curves. This poses an important problem as the simple RC model relies heavily on OCV and the only available observation of the actual process is through a single measurement of the terminal voltage. Therefore, errors in the OCV curve, in combination with the potential parameter errors in the rather crude ECM, has a large impact on the SOC estimate. The model error is further confirmed at the end of the discharge test, where the filters produce an improper overconfident estimate.

A123 Cell Tests

Moving on to the tests performed on the A123 battery cell chemistry, with estimation results shown in Figure 5.2. As with the E2 tests, all filters are initialised with an offset in SOC at 80 %. In this case, however, all filters are able to move their respective SOC estimates toward the true value. The error plots show exceptional filtering performance for the EnKF and PF, until the true SOC reaches 10 %, at which point we observe a behaviour similar to that in the E2 tests, namely that the estimate diverges from the true state. The EnKF has a final estimation error of 2.5 % while the PF finishes as 1.5 %. The UKF, however, clearly diverges at the time of 400 min. At this point the true SOC is about 35 %, which is where the OCV curve, cf. Figure 4.6b, goes from being almost completely flat to having a non-zero gradient during discharge. Relating to the E2 tests, the residual estimation scores, K_{res}, are lower in the A123 cell type, with a value of 2 for all filters except the PF.

One of the reasons for the deviating behaviour of the UKF is likely related to the filter confidence in the terminal voltage measurements. As seen in Table 5.2, the output prediction error of the UKF, in terms of RMSE, is much smaller than for the EnKF and the PF. Even though all filters are tuned identically, using the same process and measurement noise matrices, the UKF is apparently producing an SOC estimate that better conforms to the observed terminal voltage. This could imply that the OCV influences the UKF estimate more than in the other filters and that conversely the EnKF and the PF rely more on the model prediction of the SOC. The latter attitude is apparently to an advantage in the A123 cell, which has a very flat OCV curve in the range 35–90 %, as seen in Figure 4.6b. This means that voltage measurements do not provide much information about the SOC while in this region since a particular OCV value can correspond to a wide range of SOC values. Also, comparing the output prediction errors of the A123 and E2 tests, the RMSE is much higher for the A123 cell type; implying a sensitivity in the measurement function and the OCV–SOC relationship.

Joint Parameter Estimation

To handle variations in battery cell dynamics during the discharge test, we have implemented a parameter estimation approach for adaptation of the rather simple ECM used for output predictions. The results of these tests are given in the same format as the state-only filters in Figures 5.1 and 5.2, and Tables 5.1 and 5.2. The evolution of parameter estimates is plotted in Figures 5.4 and 5.5 for the two different battery cell chemistries. A total internal resistance approximation is presented in Figure 5.6 as the sum of the two resistive parameters, R_0 and R_1.

Slight improvements are observable only for the EnKF in the E2 tests: all error statistics drop and the estimation metric K_{est} increases from 3.75 to 3.90, as seen in Table 5.1. In the A123 tests, all joint versions show small improvements in estimation error statistics, perhaps best illustrated by the small decrease in MAE seen in Table 5.2. Due to the stochastic nature of the EnKF and PF, however, these small performance enhancements are difficult to verify, and it is thus appropriate to have a discussion about some of the drawbacks of this approach.

The joint filter versions are realised by augmentation of the two-dimensional state vector to also include the three ECM parameters: R_0, R_1, and C_1. As the motion model for the parameter evolution is implemented as a stochastic walk, cf. Section 4.1.2, the joint estimation scheme introduces three additional noise processes to tune. This poses a rather intangible problem, as the ECM parameters does not have a distinct physical correspondence. Ultimately, the parameter evolution is impossible to evaluate other than observable improvements in the SOC estimation sequence. Since these are of a rather small magnitude, and only for the EnKF, our investigation leaves research question (ii) unresolved.

Parameter Evolution in the Different Filters

Although it is not possible to directly evaluate or verify the parameter estimates, results from the three different filters can be compared for consistency. Similar evolution patterns in different filters could indicate that the filter innovation from the voltage measurements is enough to move the parameter estimates in a reasonable direction. For the E2 chemistry, see Figure 5.4, there's a distinguishable similarity between the UKF and EnKF R_0 estimates: the evolution curves having a similar shape and value range. The PF R_0 estimate is different in that a higher value is estimated in the middle of the estimation sequence, while also showing a quite different evolution curve. The other parameters on the other hand show very different behaviour as the cell is being discharged. In particular, the UKF C_1 estimate is almost completely static, as compared to the other two filters, which could contribute to the slightly worse SOC estimate produced by the UKF.

A similar discussion can be made for the A123 parameter estimates, which are plotted in Figure 5.5. As for the E2 parameters, the R_0 estimate is very similar when comparing UKF and EnKF, while the PF shows a rather different behaviour, especially in the final value, which is much smaller than in the other two algorithms. Figure 5.5 further displays a similar behaviour of the R_1 estimate between the three filters, but in rather different value ranges. Again, the C_1 parameter is relatively static in the UKF, and does not present any similar behaviour when comparing the EnKF with the PF.

In conclusion, the instantaneous polarisation resistance, R_0, is the only parameter

showing a distinct evolution of its estimate during the discharge test. As this is the most tangible of the ECM parameters, representing the battery cell's internal ohmic resistance, these findings are not too surprising. In fact, a similar behaviour has been observed in [55], where a second order RC model is used to estimate different resistive components of a lead-acid battery cell. In particular, the total internal resistance is examined in [55], which shows an increase as the SOC of the cell approaches zero percent. A similar behaviour can be observed in Figure 5.6 for the sum of R_0 and R_1. Note though that such an analysis of the total resistance should really be made for non-dynamic discharge tests, to only consider static impedance components. The dynamic tests performed in this study does also excite the dynamic impedance modelled by the capacitance C_1.

Comparison of Cell Chemistry Tests

In the discussions above there is a clear difference in the estimation performance of SOC when comparing the two different cell chemistry tests. Not only visible in Figures 5.1 and 5.2, but also verified by the performance metrics in Tables 5.1 and 5.2. For instance, the SOC RMSE in the E2 cell was found to be in the range 1.07–2.13 % for the state-only filters, and 0.95–2.43 % for the joint filters. In contrast, the SOC RMSE in the A123 cell was only 0.44–1.43 % for the state-only filters, and 0.41–1.28 % for the joint versions, almost half the RMSE of the E2 cell.

A possible inconsistency in the filter algorithms can be analysed by inspection of the filter estimated uncertainty. Comparing the average square-root of $\mathbf{P}_{k|k}$, which we call the filter STD, with the estimation RMSE, a consistency was observed in the uncertainty to estimation error ratio. The ratio for both versions of the UKF was 0.30–0.35 in both cell types, and 0.45–0.56 for both EnKF versions. The PF showed to have the best ratio of 0.61–0.7 in the A123 cell type, and conversely the worst ratio of 0.14–0.18 in the E2 type.

There are a number of possible reasons for the quite apparent difference in performance, detailed in the chemistry-specific discussions. The relatively bad performance in the E2 cell is likely a combination of a problematic OCV relationship to the SOC, the ECM battery model being over-simplified for the particular situation, model parameters being erroneously identified, and improper tuning of the filter noise. In connection to research question (i), it is likely that the choice of battery model should be made with regard to the battery cell chemistry type of interest. A corollary to this hypothesis is that the first order RC model is suitable for the A123 cell type.

Results in Comparison to Other Research

With the goal to investigate the validity of the EnKF in the SOC estimation problem, it is of interest to also evaluate its performance in relation to findings in other research. In the two-part paper series [24], [25], a remarkably similar investigation to the one presented in this report is performed using a sigma-point method, for state estimation only [24] as well as joint estimation of states and battery model parameters [25].

The implemented filter algorithm, referred to as CDKF, uses a first order RC ECM for terminal voltage prediction in a $LiMn_2O_4$ cell. In addition to estimating the SOC and the voltage diffusion effect, the OCV voltage hysteresis is also estimated, cf. Figure 4.6. The same UDDS discharge cycle test, described in Section 4.3.1, is performed at the constant temperature of 25 °C. Under these premises, results presented in [25] showed a SOC

estimation RMSE of 1.44 % for the CDKF and 1.18 % for the joint CDKF, when the filters were initialised with a SOC offset of 20 %, as in our own tests.

In comparison, our UKF is the closest relative among the three filters we have studied to the CDKF algorithm in [25]. Concentrating on the A123 cell type, the estimation RMSEs of our UKF were 1.43 %, in state-only estimation, and 1.28 % in the joint UKF. This is well in line with the the findings in [25], with the exception that our joint UKF performs slightly worse than the joint CDKF.

Our other filter algorithms—the EnKF, PF, and their respective joint versions—show better results than the CDKFs in [25]. The RMSE was for the EnKF 0.60 %, for the joint EnKF 0.57 %; for the PF it was 0.44 % and for the joint PF 0.41 %. Connecting to research question (iii) in Section 1.3, we interpret this as an indication that the EnKF has good potential as a nonlinear filtering algorithm in the SOC estimation problem.

6.2 EnKF Dimensions

The primary motivation for this thesis was to investigate the EnKF as a statistical filtering algorithm for estimation of SOC in battery cells. The formulation of the recursive filtering algorithm in Section 2.4.2 was succeeded by a short review of some of the important advantages and disadvantages with the ensemble idea. As most significant issues are coupled with a too small ensemble size, N, an investigation of the EnKF performance for different ensemble sizes was made, to answer research question (iv). Results of this analysis are shown in Figure 5.3, for the state-only estimation approach in the A123 cell type.

The primary problem encountered when implementing the EnKF was that the filter uncertainty was underestimated. Even though the state-space is of low dimension, a quite large number of ensemble members were required in order to obtain an acceptable uncertainty estimate, as related to the estimation error. In Figure 5.3a, the SOC RMSE is plotted for different ensemble sizes, together with the filter estimated uncertainty. It is apparent that the uncertainty is underestimated for sizes $N < 200$ but corresponds better to the actual estimation error for $N \geq 500$.

In Figure 5.3b, the uncertainty error is plotted versus ensemble size. The uncertainty error is taken to be the percentage of the total length of the estimation sequence, where the estimation error is larger than three STDs of the filter uncertainty. As expected, the uncertainty error decreases with increasing ensemble size, and shows that for the $N = 2000$ setting, the uncertainty is underestimated for only 10 % of the total sequence.

Figure 5.3c shows the performance metric K_{est}, see Section 4.5.1 for definition, as a function of ensemble size. Apparently, the estimation accuracy is not improved in the K_{est} sense for ensemble sizes above $N = 1000$ members. Finally, in Figure 5.3d, the total computation time for the complete estimation sequence is plotted. As might be expected, it increases linearly with the ensemble size.

As this thesis does not aim to design an algorithm towards any specific performance requirements, as related to accuracy and time, it does not make sense to reason about the optimal size of the ensemble in the EnKF. Although there seems to exist a definite requirement on the choice of N for which the EnKF works, as seen in Figure 5.3a, the discussion about the trade-off between accuracy and efficiency is left for future research and more specific implementations.

6.3 Methods and Simplifications

The implications of the results discussed above are of course dependent on the methodology adopted in this work. Below follows a discussion regarding some of the most important assumptions and simplifications made in this study.

Battery Modelling

Modelling of the battery dynamics is one of the parts in this thesis that greatly influenced the SOC estimation results. The choice of using the first order RC model for the cell voltage predictions is a result of evaluating the trade-off between complexity and accuracy in battery modelling. It is to be noted that higher accuracy could be obtained by using other battery modelling techniques, such as the physics-based approach, but with higher complexity and computational cost. The first order RC model is a simple equivalent electrical circuit, modelling the battery cell's input–output dynamics, and can as such not represent the underlying complex electrochemical processes.

The equivalent circuit parameters do not hold any physical significance and the parametric quantitative results are only numeric and do not quantify any physical interpretations. Lack of physical interpretations for the parameters poses disadvantages in verification of correctness of estimated parameters.

Temperature Simplifications

Battery packs in vehicles are generally temperature controlled in the range 15–35 °C, for better safety, performance, and battery life [53], [54]. The dynamic test data recorded at 25 °C was used to carry out the evaluation of the filter algorithms as this temperature represents the operating temperature of the batteries. The results thus obtained in this study are representative for 25 °C and in the mentioned temperature range. The algorithm implementation may not be accurate outside this temperature window, as the OCV hysteresis outside this range is too large to approximate with an average hysteresis curve.

Sensor Accuracy

The test data used in the analysis of the filtering algorithms was collected in a controlled lab environment, using high-quality sensor equipment and without disturbances. This means that the performance of the different SOC estimation schemes was analysed in a very fair environment setting with a low uncertainty in the measurement data. However, in real-time scenarios and EV applications, the estimation algorithm would likely need to deal with low-cost sensor equipment and unpredictable disturbances of the measurements. This poses a different setting for the posed problem, and would result in less robust, and possibly less accurate, SOC estimates as compared to what is presented in this study.

Current as Input

Due to the high accuracy of the measurement data, the observations of the cell load current were possible to use directly as an input to the system in the formulation of the estimation

problem. The alternative is to associate the current measurements with an uncertainty, as for the terminal voltage, and filter the load current as a state in the battery model.

Depending on the sensor quality in a particular EV application, the latter approach might be necessary to mitigate the effect of spurious energy demands from the battery. However, to filter the cell load current, a model is also needed for predictions of the energy demand. Difficult as it may be to make informed predictions about the energy demand in an EV, the simplest and most versatile approach would be to consider a random walk for the evolution of the load current.

Filtering of the load current also introduces a small delay in the observed energy demand. This directly affects the SOC estimate, as it is modelled according to the "coulomb counting" principle.

Filter Tuning for Comparison

The primary purpose of the thesis is to evaluate the performance of the EnKF algorithm, and the process and measurement noise parameters were first adjusted to improve the EnKF SOC estimate. The same noise matrices were then used for both the UKF and PF in order to make a fair comparison among the three algorithms, with respect to the performance of the EnKF. As each algorithm has its own advantages and disadvantages, dedicated tuning of the noise matrices might be required. It should be noted that this common setting is meant to provide a consistent platform for comparison of the three filters and may not, as such, represent the best settings for the UKF and PF.

Battery Chemistry

The two battery types investigated in this thesis are referred to as E2 and A123, which are just names for the purpose of identification in the data set. There is no certain way to identify the actual battery chemistries related to these two battery cell types, making it difficult to associate the results obtained in this report to the actual batteries used in electric vehicles. This could create a gap between the research work carried out in this thesis and the requirement in the industry.

7
Conclusion

The purpose of this thesis is to investigate the potential of the EnKF as a nonlinear statistical filtering algorithm for the problem of efficiently and accurately estimating state-of-charge in lithium-ion battery cells. Towards this end, the study at hand has presented solutions for the sub-tasks of modelling electrochemical battery cell dynamics; identifying equivalent-circuit model parameters using current and voltage measurements; formulating the estimation problem within the Bayesian inference framework; and implementing a UKF, an EnKF, and a bootstrap PF for joint estimation of immeasurable states and non-stationary parameters in lithium-ion battery cells.

Results have been presented and discussed for the obtained SOC estimation errors using the three respective filters, with and without parameter estimation, for two different battery cell chemistries. The EnKF performance has been evaluated and validated internally within this study, by comparing with the UKF and PF, as well as externally by connecting to a similar study, implementing a CDKF in a nearly identical test scenario. Furthermore, the performance of the EnKF has been investigated as a function of the ensemble size.

Major Findings

This study has shown that the EnKF is able to produce accurate SOC estimates—better than the UKF and comparable to the PF—both with and without estimation of battery model parameters. The results have furthermore shown that, compared with an external study, the EnKF is capable of producing more accurate SOC estimates than the sigma-point method CDKF.

The research has also shown that there is a lower bound for the ensemble size for which the EnKF works. Above this limit, there is a trade-off between computational cost and estimation accuracy that can be analysed for specific operation conditions in a real-time application.

The investigation of joint parameter estimation has shown that the accuracy of the SOC estimate can be sensitive to both the initial system identification and tuning of the stochastic parameter evolution. The chemistry-specific tests also suggest the battery cell type can have a large impact on the robustness of the estimation scheme, and should affect the choice of battery model, as nonlinearities in open-circuit voltage can be more or less significant.

Ideas for Future Work

Below is a list of ideas, and suggestions for extensions, of this thesis, that has emerged during the progression of this project, but had to be omitted due to either being outside the scope or due to lack of time and resources.

Noise adaptation. With the introduction of parameter estimation, the number of estimated quantities increased from two to five. Even though this does not, in general, represent a large system, the abstract nature of the ECM parameters resulted in an intangible tuning problem. A few approaches to adaptation of the process and measurement noise matrices were therefore explored but never implemented, due to time limitations. With efficiency in mind, we propose an innovation-based adaptive policy for the process noise in the three different filters, with the possibility to increase the versatility of our estimation scheme.

Noise adaptation for this particular problem has been examined extensively for EKF implementations [20], [21], [56], as noted in the related work, with indications of improved error statistics, using the adaptation schemes such as Sage–Husa [57] and covariance matching [58]. An innovation-based adaptive implementation of the UKF, using covariance matching, has also been explored in [27]. Two additional papers provide quite tangible reviews on different methods for noise adaptation in localisation and positioning problems [59], [60].

Parameter noise tuning. An alternative to the proposed solution of *noise adaptation* for the battery model parameter, a more systematic tuning process can be developed, optimising the parameter noise levels with the objective to minimise the SOC estimation error. Although this particular approach could potentially result in over-fitting tuning parameters to a particular data set, we would argue to spend more time on tuning the filters to be able to potentially find an answer to research question (ii).

Battery model extensions. Comparison of estimation results in the two different battery cell chemistry types suggested a sensitivity to the OCV–SOC relationship, particularly related to the charge–discharge hysteresis. As such, the state-vector might be extended to also include the level of hysteresis in the OCV at the present SOC level. This idea was investigated thoroughly in the two paper series [17]–[19] and [24], [25]. We believe this approach also has the potential to increase the versatility of our SOC estimation scheme and would be worth investigating while also considering the added computational cost.

Electrochemical battery models. The most significant advantage of the EnKF algorithm is the ability to handle extremely high dimensional problems. Since this thesis employed a fairly simple equivalent circuit battery model, described by only two states, the usefulness of the EnKF for the SOC estimation problem could not be evaluated completely. In contrast to the ECMs, the electrochemical battery model category requires a higher dimensionality of the state. An interesting subject for future studies is thus to investigate the EnKF in conjunction with a high-dimensional electrochemical battery model.

High dynamic test scenario. In this thesis, the SOC estimation scheme has been evaluated in one single test scenario: the UDDS test. In this test, the load current is designed to be representative for urban driving, which means that the demanded current is retained within moderate bounds. In subsequent studies, it is of course important to examine how the estimation scheme handles a higher dynamical behaviour of the demanded current, representative for, e.g., highway driving. Such a scenario would

also have implications for the temperature of the battery pack, which introduces two important problems: estimation of cell temperature and how to handle temperature variations.

Temperature estimation. Due to limited access to appropriate data sets, the problems in SOC estimation related to temperature had to be neglected. In a real application, it is unreasonable to assume a constant and homogeneous temperature throughout the battery pack, motivating a separate estimation scheme for inferring the inner battery cell temperature. With the goal to perform real-time SOC estimation in EVs, this is an important problem to consider in future research.

Temperature variations. Due to time limitations, the estimation tests had to be restricted to 25 °C, to be able to provide a meaningful discussion about other aspects of SOC estimation. It is however of recommended to in future research examine how the estimation performance is affected by the cell temperature, in particular temperature variations during discharge.

Bibliography

[1] G. L. Plett, *Battery Management Systems, Volume I: Battery Modeling*. Artech House Publishers, 2015.

[2] S. Yuan, H. Wu, and C. Yin, "State of charge estimation using the extended kalman filter for battery management systems based on the arx battery model", *Energies*, vol. 6, no. 1, pp. 444–470, 2013.

[3] R. E. Kalman, "A new approach to linear filtering and prediction problems", *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, Mar. 1960.

[4] M. H. Westbrook, *The Electric Car: Development and Future of Battery, Hybrid and Fuel-Cell Cars*, ser. IET Power and Enery Series 38. London: IET, 2001.

[5] G. L. Plett, *Battery Management Systems, Volume II: Equivalent-Circuit Methods*. Artech House Publishers, 2015.

[6] Z. Cai, G. Liu, and J. Luo, "Research state of charge estimation tactics of nickel-hydrogen battery", in *Proceedings - 2010 International Symposium on Intelligence Information Processing and Trusted Computing, IPTC 2010*, IEEE, Oct. 2010, pp. 184–187.

[7] S. Piller, M. Perrin, and A. Jossen, "Methods for state-of-charge determination and their applications", *Journal of Power Sources*, vol. 96, no. 1, pp. 113–120, 2001.

[8] A. Fotouhi, D. J. Auger, K. Propp, S. Longo, and M. Wild, "A review on electric vehicle battery modelling: From lithium-ion toward lithium–sulphur", *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 1008–1021, 2016.

[9] A. Seaman, T.-S. Dao, and J. McPhee, "A survey of mathematics-based equivalent-circuit and electrochemical battery models for hybrid and electric vehicle simulation", *Journal of Power Sources*, vol. 256, pp. 410–423, 2014.

[10] X. Hu, S. Li, and H. Peng, "A comparative study of equivalent circuit models for li-ion batteries", *Journal of Power Sources*, vol. 198, pp. 359–367, 2012.

[11] S. Nejad, D. Gladwin, and D. Stone, "A systematic review of lumped-parameter equivalent circuit models for real-time estimation of lithium-ion battery states", *Journal of Power Sources*, vol. 316, pp. 183–196, 2016.

[12] C. Lin, A. Tang, and J. Xing, "Evaluation of electrochemical models based battery state-of-charge estimation approaches for electric vehicles", *Applied Energy*, 2017.

[13] C. Lin, J. Xing, and A. Tang, "Lithium-ion battery state of charge/state of health estimation using smo for evs", *Energy Procedia*, vol. 105, pp. 4383–4388, 2017.

[14] A. Bartlett, J. Marcicki, S. Onori, G. Rizzoni, X. G. Yang, and T. Miller, "Electrochemical model-based state of charge and capacity estimation for a composite electrode lithium-ion battery", *IEEE Transactions on Control Systems Technology*, vol. 24, no. 2, pp. 384–399, 2016.

[15] R. Xiong, H. He, H. Guo, and Y. Ding, "Modeling for lithium-ion battery used in electric vehicles", *Procedia Engineering*, vol. 15, pp. 2869–2874, 2011.

[16] E. A. Wan and R. Van Der Merwe, "The unscented kalman filter for nonlinear estimation", in *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, IEEE, 2000, pp. 153–158.

[17] G. L. Plett, "Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 1. Background", *Journal of Power Sources*, vol. 134, no. 2, pp. 252–261, 2004.

[18] ——, "Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 2. Modeling and identification", *Journal of Power Sources*, vol. 134, no. 2, pp. 262–276, 2004.

[19] ——, "Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 3. State and parameter estimation", *Journal of Power Sources*, vol. 134, no. 2, pp. 277–292, 2004.

[20] Hongwen He, Rui Xiong, Xiaowei Zhang, Fengchun Sun, and JinXin Fan, "State-of-Charge Estimation of the Lithium-Ion Battery Using an Adaptive Extended Kalman Filter Based on an Improved Thevenin Model", *IEEE Transactions on Vehicular Technology*, vol. 60, no. 4, pp. 1461–1469, May 2011.

[21] R. Xiong, H. He, F. Sun, and K. Zhao, "Evaluation on State of Charge Estimation of Batteries With Adaptive Extended Kalman Filter by Experiment Approach", *IEEE Transactions on Vehicular Technology*, vol. 62, no. 1, pp. 108–117, Jan. 2013.

[22] H. He, R. Xiong, and H. Guo, "Online estimation of model parameters and state-of-charge of lifepo 4 batteries in electric vehicles", *Applied Energy*, vol. 89, no. 1, pp. 413–420, 2012.

[23] R. Xiong, H. He, F. Sun, X. Liu, and Z. Liu, "Model-based state of charge and peak power capability joint estimation of lithium-ion battery in plug-in hybrid electric vehicles", *Journal of power sources*, vol. 229, pp. 159–169, 2013.

[24] G. L. Plett, "Sigma-point Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 1: Introduction and state estimation", *Journal of Power Sources*, vol. 161, no. 2, pp. 1356–1368, 2006.

[25] ——, "Sigma-point Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 2: Simultaneous state and parameter estimation", *Journal of Power Sources*, vol. 161, no. 2, pp. 1369–1384, 2006.

[26] S. Santhanagopalan and R. E. White, "State of charge estimation using an unscented filter for high power lithium ion cells", *International Journal of Energy Research*, vol. 34, no. 2, pp. 152–163, Feb. 2010.

[27] F. Sun, X. Hu, Y. Zou, and S. Li, "Adaptive unscented Kalman filtering for state of charge estimation of a lithium-ion battery for electric vehicles", *Energy*, vol. 36, no. 5, pp. 3531–3540, 2011.

[28] W. He, N. Williard, C. Chen, and M. Pecht, "State of charge estimation for electric vehicle batteries using unscented kalman filtering", *Microelectronics Reliability*, vol. 53, no. 6, pp. 840–847, 2013.

[29] S. Schwunk, N. Armbruster, S. Straub, J. Kehl, and M. Vetter, "Particle filter for state of charge and state of health estimation for lithium–iron phosphate batteries", *Journal of Power Sources*, vol. 239, pp. 705–710, 2013.

[30] X. Liu, Z. Chen, C. Zhang, and J. Wu, "A novel temperature-compensated model for power Li-ion batteries with dual-particle-filter state of charge estimation", *Applied Energy*, vol. 123, pp. 263–272, 2014.

[31] Y. Wang, C. Zhang, and Z. Chen, "A method for state-of-charge estimation of LiFePO4 batteries at dynamic currents and temperatures using particle filter", *Journal of Power Sources*, vol. 279, pp. 306–311, 2015.

[32] M. Roth, G. Hendeby, C. Fritsche, and F. Gustafsson, "The ensemble kalman filter: A signal processing perspective", *arXiv preprint arXiv:1702.08061*, 2017.

[33] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics", *Journal of Geophysical Research*, vol. 99, no. C5, p. 10 143, 1994.

[34] G. Burgers, P. Jan van Leeuwen, and G. Evensen, "Analysis Scheme in the Ensemble Kalman Filter", *Monthly Weather Review*, vol. 126, no. 6, pp. 1719–1724, Jun. 1998.

[35] J. Mandel, J. Beezley, J. Coen, and M. Kim, "Data assimilation for wildland fires", *IEEE Control Systems Magazine*, vol. 29, no. 3, pp. 47–65, Jun. 2009.

[36] J. Anderson, "Ensemble Kalman filters for large geophysical applications", *IEEE Control Systems Magazine*, vol. 29, no. 3, pp. 66–82, Jun. 2009.

[37] G. Evensen, "The ensemble Kalman filter for combined state and parameter estimation", *IEEE Control Systems Magazine*, vol. 29, no. 3, pp. 83–104, Jun. 2009.

[38] S. Lakshmivarahan and D. Stensrud, "Ensemble Kalman filter", *IEEE Control Systems Magazine*, vol. 29, no. 3, pp. 34–46, Jun. 2009.

[39] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.

[40] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, New Jersey: Prentice-Hall, 1979.

[41] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems", I. Kadar, Ed., International Society for Optics and Photonics, Jul. 1997, p. 182.

[42] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation", *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, Nov. 2004.

[43] M. Rhudy, Y. Gu, J. Gross, and M. R. Napolitano, "Evaluation of matrix square root operations for ukf within a uav gps/ins sensor fusion application", *International Journal of Navigation and Observation*, vol. 2011, 2012.

[44] N. J. Gordon, D. J. Salmond, and A. F. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation", in *IEE Proceedings F (Radar and Signal Processing)*, IET, vol. 140, 1993, pp. 107–113.

[45] C. Campestrini, M. F. Horsche, I. Zilberman, T. Heil, T. Zimmermann, and A. Jossen, "Validation and benchmark methods for battery management system functionalities: State of charge estimation algorithms", *Journal of Energy Storage*, vol. 7, pp. 38–51, 2016.

[46] C. Campestrini, T. Heil, S. Kosch, and A. Jossen, "A comparative study and review of different Kalman filters by applying an enhanced validation method", *Journal of Energy Storage*, vol. 8, pp. 142–159, 2016.

[47] P. L. Houtekamer, H. L. Mitchell, P. L. Houtekamer, and H. L. Mitchell, "Data Assimilation Using an Ensemble Kalman Filter Technique", *Monthly Weather Review*, vol. 126, no. 3, pp. 796–811, Mar. 1998.

[48] P. Houtekamer and H. L. Mitchell, "Ensemble Kalman filtering", *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 613, pp. 3269–3289, Oct. 2005. [Online]. Available: `http://doi.wiley.com/10.1256/qj.05.135`.

[49] T. M. Hamill, "Ensemble-based atmospheric data assimilation", *Predictability of weather and climate*, pp. 124–156, 2006.

[50] F. Le Gland, V. Monbet, and V. Tran, "Large sample asymptotics for the ensemble Kalman filter", in *The Oxford Handbook of Nonlinear Filtering*, D. Crisan and B. Rozovskii, Eds. Oxford University Press, 2011, pp. 598–634.

[51] L. Ljung and T. Glad, *Modeling of Dynamic Systems*. Englewood Cliffs, New Jersey: Prentice Hall, 1994.

[52] G. L. Plett, *ESC Model Toolbox*, 2015. [Online]. Available: `http://mocha-java.uccs.edu/BMS1/CH02/ESCtoolbox.zip` (visited on 05/09/2017).

[53] A. Pesaran, M. Keyser, G.-H. Kim, S. Santhanagopalan, and K. Smith, "Tools for designing thermal management of batteries in electric drive vehicles", in *Large Lithium Ion Battery Technology & Application Symposia*, Advanced Automotive Battery Conference (Pasadena, CA, Feb. 4–8, 2013), NREL, 2013.

[54] Q. Wang, B. Jiang, B. Li, and Y. Yan, "A critical review of thermal management models and solutions of lithium-ion batteries for the development of pure electric vehicles", *Renewable and Sustainable Energy Reviews*, vol. 64, pp. 106–128, Oct. 2016.

[55] "A method to monitor valve-regulated lead acid cells", *Journal of Power Sources*, vol. 74, no. 2, pp. 234–239, 1998.

[56] J. Han, D. Kim, and M. Sunwoo, "State-of-charge estimation of lead-acid batteries using an adaptive extended Kalman filter", *Journal of Power Sources*, vol. 188, no. 2, pp. 606–612, 2009.

[57] A. P. Sage and G. W. Husa, "Adaptive filtering with unknown prior statistics", in *Joint Automatic Control Conference*, 1969, pp. 760–769.

[58] R. Mehra, "Approaches to adaptive filtering", *IEEE Transactions on automatic control*, vol. 17, no. 5, pp. 693–698, 1972.

[59] A. H. Mohamed and K. P. Schwarz, "Adaptive Kalman Filtering for INS/GPS", *Journal of Geodesy*, vol. 73, no. 4, pp. 193–203, May 1999.

[60] D. Jwo and S. Chang, “Particle swarm optimization for GPS navigation Kalman filter adaptation”, *Aircraft Engineering and Aerospace Technology*, vol. 81, no. 4, pp. 343–352, Jul. 2009.