| | |
|---|---|
| **Titre:** Title: | Distributionally Robust Extended Kalman Filtering for State-of-Charge Estimation of Lithium-ion Batteries |
| **Auteur:** Author: | Donald Fisher |
| **Date:** | 2026 |
| **Type:** | Mémoire ou thèse / Dissertation or Thesis |
| **Référence:** Citation: | Fisher, D. (2026). Distributionally Robust Extended Kalman Filtering for State-of-Charge Estimation of Lithium-ion Batteries [Master's thesis, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/76507/ |



| | |
|---|---|
| **URL de PolyPublie:** PolyPublie URL: | https://publications.polymtl.ca/76507/ |
| **Directeurs de recherche:** Advisors: | Keyhan Sheshyekani, & Antoine Lesage-Landry |
| **Programme:** Program: | Génie électrique |

**POLYTECHNIQUE MONTRÉAL**
affiliée à l'Université de Montréal

# Distributionally Robust Extended Kalman Filtering for State-of-Charge Estimation of Lithium-ion Batteries

**DONALD FISHER**
Département de génie électrique

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
Génie électrique

Mai 2026

**POLYTECHNIQUE MONTRÉAL**
affiliée à l'Université de Montréal

Ce mémoire intitulé :

**Distributionally Robust Extended Kalman Filtering for State-of-Charge Estimation of Lithium-ion Batteries**

présenté par **Donald FISHER**
en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
a été dûment accepté par le jury d'examen constitué de :

**Jean MAHSEREDJIAN**, président
**Keyhan SHESHYEKANI**, membre et directeur de recherche
**Antoine LESAGE-LANDRY**, membre et codirecteur de recherche
**Amin GHAZANFARI**, membre

# DEDICATION

*If only we could know when we were in the good old days before we left them.*
*Someone should write a song about that.*

## ACKNOWLEDGEMENTS

I would like to express my deep gratitude to all the people who contributed, directly or indirectly, to the completion of this Master's thesis.

Firstly, I would like to thank Mitacs Accelerate, Serafin Solutions, and Prof. Keyhan Sheshyekani for funding the project and putting forth the idea, and trusting in me, which made this project possible.

I give my eternal thanks to my research supervisor, Prof. Antoine Lesage-Landry, for his guidance, dedication and scientific rigour throughout the project. His assistance has been essential at every stage of my journey.

To Mahmoud, Eyerusalem and Hoa, for their feedback, critiques and ideas in our team meetings.

To my friends at LORER for their camaraderie and collaboration before exams, at the pub, or at all-you-can-eat sushi. To my friends outside of Polytechnique, for indulging my musings and providing enriching experiences outside of the lab.

To my family and Sam, for their support and encouragement in pursuing further studies, and for being an infinite source of comfort during challenging times. Thank you for being my greatest supporters.

To all those who recognize themselves here, thank you from the bottom of my heart.

# RÉSUMÉ

Les batteries lithium-ion sont présentement le choix principal pour le stockage d'énergie dans les véhicules électriques (*electric vehicle*, EV). Il est donc essentiel d'avoir une estimation précise et fiable de l'état de charge (*state of charge*, SOC) en temps réel afin de prévoir l'autonomie de la batterie d'une manière fiable et d'éviter sa dégradation due à une surcharge ou à une décharge profonde. Les méthodes d'estimation actuelles basées sur des modèles de circuits équivalents et des filtres, telles que le filtre de Kalman étendu (*extended Kalman filter*, EKF) et le filtre de Kalman sans parfum (*unscented Kalman filter*, UKF), fournissent une estimation efficace du SOC en temps réel. Cependant, elles supposent généralement un bruit de processus et un bruit de mesure gaussien, avec une moyenne et une covariance connues, ce qui peut ne pas être le cas dans la pratique. Le filtre de Kalman robuste en distribution utilise la distance de Wasserstein pour construire un ensemble d'ambiguïté autour d'un bruit de processus et de mesure nominal, se prémunissant ainsi contre toutes les déviations dans les valeurs nominales. Dans ce travail, nous proposons un nouveau filtre robuste en distribution de Kalman étendu (*distributionally robust extended Kalman filter*, DREKF), qui étend cette méthodologie au scénario nonlinéaire et nous l'appliquons au problème de l'estimation de l'état de la batterie. Nous comparons le DREKF contre l'EKF et l'UKF utilisant des simulations basées sur un ensemble de données ouvertes provenant d'un programme de conduite sur dynamomètre urbain. Les résultats obtenus pour un éventail de températures démontrent que le DREKF atteint une précision comparable à celle des approches EKF et UKF conventionnelles dans des conditions nominales, tout en offrant une meilleure résilience aux valeurs aberrantes des mesures.

## ABSTRACT

Lithium-ion batteries (LIBs) are the core energy storage component of electric vehicles (EVs), making real-time, accurate, and robust state-of-charge (SOC) estimation critical for reliable range prediction and preventing degradation due to overcharging or deep discharge. Current model-based estimation methods, which utilize equivalent circuit models and filters, such as the extended Kalman filter (EKF) and unscented Kalman filter (UKF), provide effective real-time SOC estimation. However, they typically assume a Gaussian process and measurement noise with known mean and covariance, which may not hold in practice. Distributionally robust Kalman filtering (DRKF) utilizes the Wasserstein distance to construct an ambiguity set around a nominal process and measurement noise, thereby hedging against any errors in these nominal assumptions. However, existing literature is only applicable in the linear setting. In this paper, we propose the distributionally robust extended Kalman filter (DREKF), which extends this methodology to the nonlinear battery state estimation problem. We benchmark the DREKF against the EKF and UKF using simulations based on an open-source urban dynamometer driving schedule dataset. Results across multiple temperatures and operating conditions demonstrate that the DREKF achieves accuracy comparable to conventional EKF and UKF approaches under nominal conditions, while offering improved resilience to measurement outliers.

# TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# LIST OF SYMBOLS AND ABBREVIATIONS

Acronyms

| | |
|---|---|
| AUKF | Adaptive Unscented Kalman Filter |
| BMS | Battery Management System |
| CC | Coulomb Counting |
| CNN | Convolutional Neural Net |
| DL | Deep Learning |
| DoD | Depth of Discharge |
| DREKF | Distributionally Robust Extended Kalman Filter |
| DRKF | Distributionally Robust Kalman Filter |
| DRO | Distributionally Robust Optimization |
| ECM | Equivalent Circuit Model |
| EKF | Extended Kalman Filter |
| EM | Electrochemical Model |
| EV | Electric vehicle |
| FRTSUKF | Fuzzy Robust Two-Stage Unscented Kalman Filter |
| GRU | Gated Recurrent Unit |
| HPPC | Hybrid Pulse Power Characterization |
| KCL | Kirchhoff's Current Law |
| KF | Kalman filter |
| KVL | Kirchhoff's Voltage Law |
| LIB | Lithium-ion Battery |
| LSTM | Long Short-Term Memory |
| MAE | Mean Absolute Error |
| ML | Machine Learning |
| MLP | Multi-Layer Perceptron |
| OCV | Open-Circuit Voltage |
| ODE | Ordinary Differential Equation |
| P2D | Pseudo-2-d |
| PDE | Partial Differential Equation |
| PF | Particle Filter |
| RLS | Recursive Least Squares |
| RMSE | Root Mean Square Error |

| | |
|---|---|
| RNN | Recurrent Neural Net |
| SDP | Semidefinite Program |
| SOC | State-of-Charge |
| SOH | State-of-Health |
| SPM | Single Particle Model |
| SRUKF | Square-Root Unscented Kalman Filter |
| SVM | Support Vector Machine |
| SW | Sliced Wasserstein |
| UDDS | Urban Dynamic Dynamometer Schedule |
| UKF | Unscented Kalman Filter |

Symbols

| | |
|---|---|
| $\alpha_j$ | Integrating factor at RC pair $j$ |
| $\mathbf{A}$ | State transition matrix |
| $\mathbf{B}$ | Control input matrix |
| $\mathcal{B}_\theta(\mathbb{Q})$ | Wasserstein ambiguity set centered at a nominal distribution, $\mathbb{Q}$ |
| $c_j$ | Capacitance value in a RC pair, at pair $j$ |
| $\chi_i$ | Sigma points at index $i$ generated by the unscented transform |
| $\mathbf{C}$ | Measurement transtion matrix |
| $\mathbf{D}$ | Feedthrough matrix |
| $\Delta t$ | Sample interval |
| $\epsilon$ | Algorithm update tolerance |
| $\mathcal{F}_k$ | The set of all admissible estimators at time $k$ |
| $f(\cdot)$ | Nonlinear state transition function |
| $\mathbf{F}_k$ | Jacobian of $f(\cdot)$ at time $k$ |
| $h(\cdot)$ | Nonlinear measurement functions |
| $\mathbf{H}_k$ | Jacobian of $h(\cdot)$ at evaluated at the prior state at time $k$ |
| $i_k$ | Current at time $k$ |
| $\mathbf{I}$ | The identity matrix |
| $j$ | Index for RC pairs |
| $k$ | Discrete time index |
| $\mathbf{K}$ | Kalman gain |
| $\mathbf{K}_k^*$ | The worst-case Kalman gain |
| $L$ | Dimension of the sigma point state vector used in the UKF |
| $\\| \cdot \\|_2$ | $\ell_2$-norm or Euclidean norm |

| | |
|---|---|
| $n_x, n_u, n_y$ | Dimensions of the state, input and measurement vectors, respectively |
| $n_p$ | Dimension of sigma vectors used in the unscented transform |
| $\otimes$ | The product measure operator |
| $\phi$ | A generic state estimator function |
| $\boldsymbol{\psi}$ | True ECM parameter vector |
| $\hat{\boldsymbol{\psi}}$ | Estimated ECM parameter vector |
| $\mathcal{P}$ | Set of probability distributions |
| $\mathcal{P}(\Omega)$ | The set of Borel probability measures with support $\Omega \subseteq \mathbb{R}^n$ |
| $\mathcal{P}_2(\Omega)$ | The subset of measures with finite second-order moments |
| $\mathbf{P}_k$ | Error covariance Matrix at time $k$ |
| $\mathbb{P}, \mathbb{Q}$ | Probability distribution |
| $Q_n$ | Nominal battery capacity (Ah) |
| $\mathbf{Q}_k$ | Process noise covariance at time $k$ |
| $\hat{\mathbb{Q}}$ | Nominal probabilty Distribution |
| $r_0$ | Instantaneous voltage drop in the ECM model |
| $r_j$ | Resistance value in a RC pair, at pair $j$ |
| $\mathbf{R}_k$ | Process noise covariance at time $k$ |
| $\mathbb{R}^n$ | Set of real numbers in dimension $n$ |
| $\mathcal{S}^n$ | The set of all symmetric matrices in $\mathbb{R}^{n\times n}$ |
| $\mathcal{S}^n_+$ | The cone of symmetric positive semidefinite matrices in $\mathcal{S}^n$ |
| $\mathcal{S}^n_{++}$ | The cone of symmetric positive definite matrices in $\mathcal{S}^n$ |
| $s_k$ | State-of-charge at time $k$ |
| $\hat{\boldsymbol{\Sigma}}^-_{\mathbf{x}_0}$ | Nominal state covariance matrix at time $k$ |
| $\hat{\boldsymbol{\Sigma}}_{\mathbf{w}_k}$ | Nominal process noise covariance matrix at time $k$ |
| $\hat{\boldsymbol{\Sigma}}_{\mathbf{v}_k}$ | Nominal measurement covariance matrix at time $k$ |
| $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k\|k-1}}$ | The pseudo-nominal prior state covariance |
| $\boldsymbol{\Sigma}_{\mathbf{x}_{k\|k-1}}$ | Prior state covariances at time $k$ |
| $\boldsymbol{\Sigma}_{\mathbf{x}_{k\|k}}$ | Posterior state covariance matrix at time $k$ |
| $\boldsymbol{\Sigma}^*_{\mathbf{x}_{k\|k-1}}$ | Worst-case prior state covariance matrix at time $k$ |
| $\boldsymbol{\Sigma}^*_{\mathbf{x}_{k\|k-1}}$ | Worst-case posterior state covariance matrix at time $k$ |
| $\boldsymbol{\Sigma}^*_{\mathbf{v}_k}$ | Worst-case measurement covariance |
| $t$ | Continuous time index (seconds) |
| $\boldsymbol{\theta}$ | Wasserstein radius vector |
| $T$ | Temperature in °C or Kelvin |
| $\tau_j$ | Time constant for RC pair $j$ |
| $u_{\mathrm{oc}}(s)$ | Open-circuit voltage |

| | |
|---|---|
| $u'_{\mathrm{oc}}(s)$ | 1st-order derivative of the OCV with respect to the state-of-charge |
| $u_j$ | Transient voltage at RC pair, $j$ |
| $\dot{u}_j$ | 1st order time derivative of voltage at $j$ RC pair |
| $u_{\mathrm{L}}$ | Load voltage |
| $\mathbf{v}_k$ | Measurement noise vector at time $k$ |
| $W_i$ | Weights applied to the respective sigma points |
| $\mathrm{W}_2(\mathbb{P}, \mathbb{Q})$ | Order-2 Wasserstein distance of two distributions $\mathbb{P}$ and $\mathbb{Q}$ with respect to the Euclidean norm |
| $\mathbf{w}_k$ | Process noise vector |
| $x_{\ell,k}$ | $\ell$-th element in $\mathbf{x}$ at timestep, $k$ |
| $\mathbf{x}_k$ | State vector at time $k$ |
| $\hat{\mathbf{x}}_{k\|k-1}$ | The prior state estimate |
| $\mathcal{Y}_i$ | Propagated measurement sigma points |
| $\hat{\mathbf{y}}_k$ | Predicted measurement Vector at time $k$ |
| $\mathbf{y}_k$ | Actual measurement vector at time $k$ |
| $\mathbf{Y}, \mathbf{Z}$ | Schur complementary variables used to convert the optimization into a convex problem |

# CHAPTER 1 INTRODUCTION

Lithium-ion batteries (LIBs) are widely used in electric vehicles (EVs). It is projected that the global EV stock will expand to 140 million by 2030 as the need to decarbonize transportation increases [2]. LIBs are the most popular choice for energy storage in EVs due to their high energy density, long lifespan, light weight, and low self-discharge rate [3] relative to other battery chemistries. LIB demand for EVs increased by about 65% to 550 GWh in 2022, from about 330 GWh in 2021, primarily as a result of growth in electric passenger car sales [4]. To ensure the battery's lifespan [5] and alleviate range anxiety [6], battery management systems (BMS) must provide accurate and robust real-time state-of-charge (SOC) estimation. Precise SOC estimation further proves valuable in simulation contexts, wherein the battery's SOC and voltage characteristics can be modelled to inform the design and specification of requirements for an EV's powertrain [7, 8].

The SOC, $s$ of a battery is defined as the ratio of the remaining capacity over the full capacity of the battery, given as:

$$s = \frac{Q_t}{Q_\mathrm{n}},$$

where $Q_t$ represents the current remaining capacity at time $t$, and $Q_n$ represents the nominal capacity given by the manufacturer. However, unlike the physical fuel tank in a combustion vehicle, an EV battery's remaining capacity, $Q_t$, cannot be measured directly. It must be inferred from external signals, namely voltage, current, and temperature.

Accurate SOC estimation remains challenging due to the inherent nonlinearity of battery dynamics and the highly variable operating environments of EVs [9]. The foundational method for tracking energy is Coulomb Counting (CC) [10]. While computationally trivial, it is an "open-loop" process. It relies on the integration of measured current over time, meaning that any sensor bias or measurement noise is integrated into a cumulative error that grows without bound. Furthermore, it depends heavily on an accurate initial SOC value, which is rarely available in dynamic operating conditions.

To correct this drift, industry standards often utilize open circuit voltage (OCV) look-up tables. By correlating the battery's voltage at rest to its charge level, the SOC can be periodically recalibrated. However, batteries in modern EVs rarely reach a true rest state for the several hours required to eliminate polarization effects. Additionally, Li-ion batteries display different OCV values, depending on whether the battery is charging or discharging, a phenomenon known as hysteresis, which adds additional complexity to determining the SOC.

To bridge the gap between current integration and direct voltage measurement, equivalent circuit models (ECMs) were developed to predict the transient behaviours of LIBs. These models represent the battery as a network of resistors and capacitors that mimic the internal electrochemical dynamics. The extended Kalman filter (EKF) and the unscented Kalman filter (UKF) have emerged as frequently implemented algorithms for SOC estimation, as they are well-suited to handle the nonlinear, time-varying, and partially unobservable nature of LIBs. From the ECM, a set of state-space relations can be formulated by combining the transient voltage equations from the ECM, CC and the OCV-SOC curves to predict the SOC and then correct it based on the measured terminal voltage in real-time.

However, these filters are built upon the assumption that the measurement noise is known and Gaussian. In real-world environments, rapid acceleration, regenerative braking, and switching operations of power electronic devices in EVs result in substantial transient current variations [11]. Current sensors and acquisition systems are often susceptible to interference from electromagnetic pulses and relay switching transients, leading to spike-voltage and current noise. Such disturbances typically exhibit non-Gaussian characteristics [11, 12]. When faced with non-Gaussian noise or significant initial estimation bias, the performance of traditional Gaussian filters degrades substantially, leading to poor estimation accuracy and convergence issues [13, 14]. This potentially shifting and non-Gaussian noise dynamics motivate the use of robust estimation frameworks. While robust estimation methods like the $H_\infty$ filter have been proposed to handle worst-case scenarios, they are sometimes seen as overly complex to implement.

In this context, the distributionally robust Kalman filter (DRKF) provides an interesting framework for addressing non-Gaussian and uncertain noise distributions for LIB SOC estimation. By considering all probability distributions within an ambiguity set, the DRKF ensure robustness against errors in the underlying noise distributions, an advantage when the true noise distribution is unknown or non-Gaussian [15]. This approach leverages convex duality to reformulate worst-case expectation problems into tractable convex problems, enabling a worst-case covariance estimate to be obtained for a recursive filter in a computationally efficient manner.

This Master's thesis applies DRKF to the LIB SOC estimation problem for EVs and, in the process, formulates the distributionally robust extended Kalman filter (DREKF). The remainder of this chapter introduces the fundamental concepts underpinning this work. Section 1.1 details the theoretical background of developing the ECM. Section 1.2 presents the filtering mechanism of the EKF and UKF. Section 1.3 then presents distributionally robust optimization and its relevance to state estimation. Finally, the main contributions, publications, and presentations associated with this Master's thesis are summarized in Section 1.4.

## 1.1 Equivalent Circuit Models

The requirement in battery modelling with respect to battery management is to have a model that allows for the observation of important states and parameters pertaining to the SOC, in this case, based on non-invasive measurements. Possible non-invasive measurements include the voltage, current and temperature on the battery surface. To this effect, AC and DC ECMs have been developed [16]. However, the AC approach is limited to laboratory analysis [16], and so for practical reasons we focus on DC ECMs for real-time SOC estimation.

As mentioned previously, there exists a monotonic relationship between a battery's terminal voltage at rest and its SOC. After sufficient rest time to eliminate transient effects, the measured OCV can be mapped to the SOC to obtain an experimentally derived OCV–SOC curve. The OCV-SOC curve is subject to hysteresis and holds differing values depending on whether the battery is charging or discharging, which are particularly pronounced at lower SOC values, as is shown in [ [1], Figure 1.1]. Here, the Depth of Discharge (DoD) is displayed, which is the inverse of SOC. It highlights that at a 0.2 DoD (80% SOC), the difference between charging and discharging OCVs is about 0.02 V. However, at 0.8 DoD (20% SOC), this difference is roughly 0.06 V. While the OCV-SOC function cannot be used directly for real-time estimation, it does provide an absolute reference for SOC that is not subject to integration drift. Thus, it can be valuable for the recalibration of the SOC, although care must be taken to identify the appropriate OCV depending on whether the battery is charging or discharging.

### 1.1.1 ECM Topologies

ECMs enable real-time SOC by providing a link between the OCV-SOC function and the measurable terminal voltage through the combination of CC, Kirchhoff's voltage law (KVL) and Kirchhoff's current law (KCL). ECMs consist of the following components to model typical LIB behaviour:

- A voltage source representing the OCV of the battery. We denote the OCV as a polynomial function of the SOC [16,17] by $u_{\text{oc}}(s) \in \mathbb{R}$, where $s$ represents the SOC.
- A resistor $r_0 > 0$, which captures the instantaneous voltage drop that occurs when a current is applied.

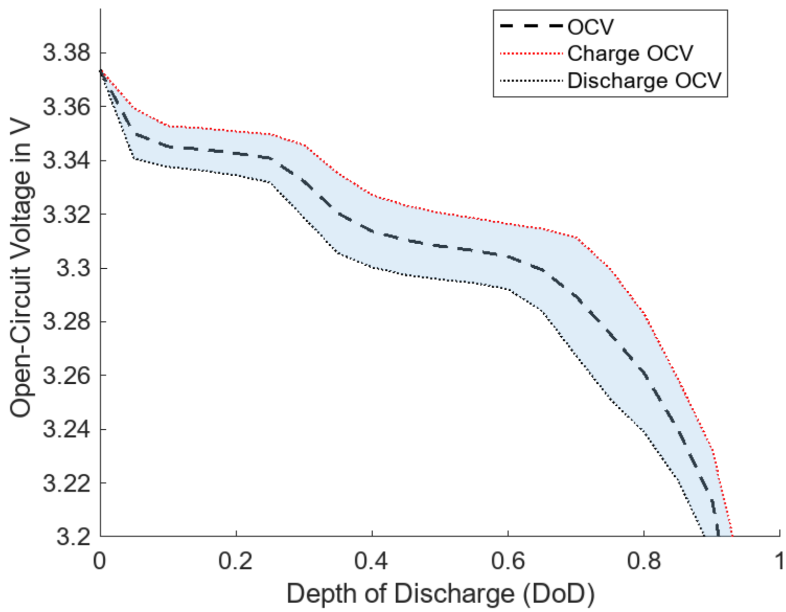



Figure 1.1 A fitted OCV-SOC curve based on empirical data for the LG18650 battery [1].

- A set of $N$ resistor-capacitor (RC) pairs, parameterized by $r_j$ and $c_j$ for $j = 1, 2, \ldots, N$, which model the voltage relaxation and polarization effects [16, 17].

ECM topologies used in literature offer a trade-off between simplicity and accuracy. The Rint model is the simplest and most basic battery model [18], consisting of a constant voltage source in series with a single resistor. It is easy to implement and computationally efficient, but cannot capture dynamic voltage behaviour or transient effects, therefore limiting its application to steady-state conditions.

The Thevenin model adds one RC pair to represent polarization dynamics, improving transient response and enabling reasonable accuracy for many control applications, though it still struggles with complex, multi-time-scale behaviour and variations as the battery ages [18,19]. The $N$-RC model extends this idea by including multiple RC branches, allowing it to capture the voltage response in different timescales [19]. While theoretically including more RC pairs can enhance the accuracy, it comes at the cost of higher computational load and potential overfitting [18]. In this Master's thesis, we choose the $N$-RC model, and set $N = 2$ (2-RC) because it yields a good level of accuracy in terms of voltage response, while only subject to limited computational complexity [16, 17]. The resulting circuit is shown in Fig. 1.2. We let $i(t) \in \mathbb{R}$ denote the applied current at time $t$, adopting the convention that discharge current is positive, and let $u_{\mathrm{L}}(t) \in \mathbb{R}$ denote the load voltage at time $t$. Also worth mentioning are the Partnership for a New Generation of Vehicle model and the Gummel-Null line model, which offer superior accuracy but at a higher computational cost due to more complex cal-

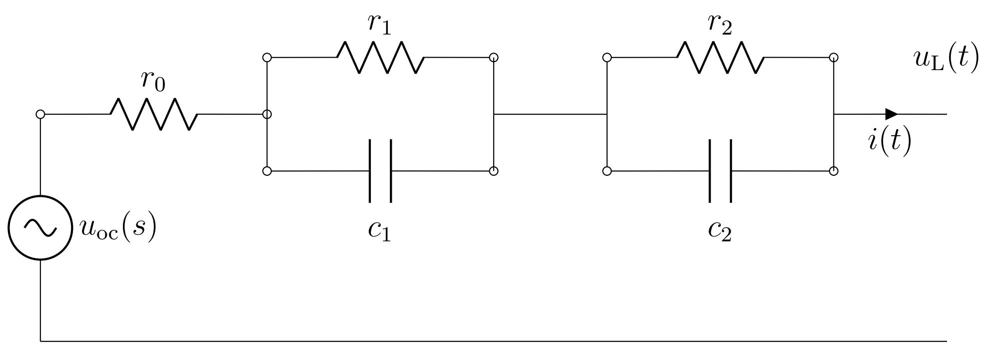



Figure 1.2 2-RC ECM

culations. They are more often used in aerospace, which demands higher precision and has a less resource-constrained management system [18]. Applying KVL and KCL to our 2-RC model, $u_{\mathrm{L}}(t)$ is given by:

$$u_{\mathrm{L}}(t) = u_{\mathrm{oc}}(s) - r_0 i(t) - u_1(t) - u_2(t), \tag{1.1}$$

where $u_1(t)$ and $u_2(t)$ represent the transient voltage drops across the first and second RC pairs, respectively. The polynomial function for $u_{\mathrm{oc}}(s)$, as well as the values for the resistors and capacitors, are typically fitted with data from hybrid pulse power characterization (HPPC) tests, conducted across a range of temperatures [16] and are briefly discussed in Section 1.1.2. The dynamics of $u_1$ and $u_2$ are governed by the following ordinary differential equations (ODEs):

$$\dot{u}_1 = \frac{1}{c_1}\left(i(t) - \frac{u_1}{r_1}\right), \tag{1.2}$$

$$\dot{u}_2 = \frac{1}{c_2}\left(i(t) - \frac{u_2}{r_2}\right), \tag{1.3}$$

where $\dot{u}_1$ and $\dot{u}_2$ denote the time derivatives of the respective RC pair voltages. To track the evolution of the SOC, we employ CC [10]. Letting $s_t$ denote the SOC at time $t$ in seconds and assuming the prior SOC $s_{t-1}$ is known, the updated SOC $s_t$ is given by:

$$s_t = s_{t-1} - \frac{1}{3600 Q_{\mathrm{n}}} \int_{t-1}^{t} i(\tau) \mathrm{d}\tau, \tag{1.4}$$

where $Q_{\mathrm{n}} > 0$ is the battery capacity in Ah, and the negative sign reflects the convention that positive current corresponds to discharge. To obtain a discretized state-space model that captures both the SOC evolution and transient voltage dynamics, we solve ODEs (1.2)–(1.3) analytically using the integrating factor

$$\alpha_j = \mathrm{e}^{-\frac{\Delta t}{\tau_j}}, \quad j = 1, 2, \ldots, N, \tag{1.5}$$

where $\Delta t$ is the sampling interval and $\tau_j = r_j c_j$ is the time constant. We index the discretized time horizon by $k \in \mathbb{N}$, corresponding to a sampling period of $\Delta t$. A process noise vector $\mathbf{w}_k \in \mathbb{R}^{3\times1}$ is included to account for current sensor noise, modelling inaccuracies, and unmodelled hysteresis effects. Combining (1.4) and the discretized solutions to (1.2)–(1.3) yields the following discrete-time state dynamics:

$$\begin{bmatrix} s_k \\ u_{1,k} \\ u_{2,k} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha_1 & 0 \\ 0 & 0 & \alpha_2 \end{bmatrix} \begin{bmatrix} s_{k-1} \\ u_{1,k-1} \\ u_{2,k-1} \end{bmatrix} + \begin{bmatrix} -\frac{\Delta t}{3600Q} \\ r_1(1-\alpha_1) \\ r_2(1-\alpha_2) \end{bmatrix} i_k + \mathbf{w}_k, \tag{1.6}$$

where $\alpha_1, \alpha_2$ are defined in (1.5). To obtain a discretized measurement equation, a scalar measurement noise term $v_k \in \mathbb{R}$ is added to (1.1), which captures voltage sensor noise and residual modelling error. Hereinafter, $u_{\mathrm{L},k} \in \mathbb{R}$ denotes the measured terminal voltage, while $\hat{u}_{\mathrm{L},k} \in \mathbb{R}$ denotes the model-based estimate of the load voltage at time step $k$. The corresponding measurement equation is:

$$\hat{u}_{\mathrm{L},k} = u_{\mathrm{oc}}(s_k) - r_0 i_k - u_{1,k} - u_{2,k} + v_k. \tag{1.7}$$

We re-express (1.6) and (1.7) compactly in the standard discrete-time stochastic state-space form and obtain

$$\begin{aligned} \mathbf{x}_k &= \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{v}_k, \end{aligned} \tag{1.8}$$

where $\mathbf{x}_k \in \mathbb{R}^{n_x}$ denotes the system state vector, $u_k \in \mathbb{R}^{n_u}$ denotes our control input $i(t)$, $\mathbf{y}_k \in \mathbb{R}^{n_y}$ denotes the measured output, and $\mathbf{v}_k$ represents a generalized measurement noise vector at time step $k$. The matrices $\mathbf{A} \in \mathbb{R}^{n_x\times n_x}$ and $\mathbf{B} \in \mathbb{R}^{n_x\times n_u}$ are directly obtained from (1.6), while $\mathbf{C} \in \mathbb{R}^{n_y\times n_x}$ and $\mathbf{D} \in \mathbb{R}^{n_y\times n_u}$ follow from the measurement equation (1.7). For the 2-RC ECM, we have dimensions $n_x = 3$, $n_u = 1$, and $n_y = 1$. Note that the measurement equation is nonlinear due to the polynomial nature of $u_{\mathrm{oc}}(s_k)$. The state vector is written component-wise as

$$\mathbf{x}_k = \begin{bmatrix} x_{1,k} \\ x_{2,k} \\ \vdots \\ x_{n_x,k} \end{bmatrix},$$

where $x_{\ell,k},\ \ell = 1, 2, \ldots N$, denotes the $\ell$-th component at time step $k$.

### 1.1.2 Parameter Identification

The reliability and accuracy of the ECM for estimating $s_k$ and $u_{\mathrm{L},k}$ depend on how precisely its parameters are identified, since each parameter directly shapes the voltage response used by the estimator. The numerical values of these parameters determine both the magnitude and dynamics of voltage transients during current changes. For example, a higher $r_0$ increases the instantaneous voltage drop under load, while a larger $\tau_j$ slows the relaxation behaviour, increasing the time taken for the load voltage to settle to the OCV value following a current pulse. If these values are inaccurate, the model produces an incorrect predicted terminal voltage $\hat{u}_{\mathrm{L},k}$, and the resulting error propagates into the SOC estimation. The ohmic resistance $r_0$, polarization resistances $r_j$, and polarization capacitances $c_j$, all vary with temperature $T$ and $s$. Accordingly, each parameter is identified as a function of these variables, e.g., $r_0(s, T)$, $r_j(s, T)$, and $c_j(s, T)$. We collect all model parameters into the vector

$$\boldsymbol{\psi} = \begin{bmatrix} r_0 \\ r_1 \\ c_1 \\ r_2 \\ c_2 \end{bmatrix} \in \mathbb{R}^5, \tag{1.9}$$

where the dependence on $s$ and $T$ is left implicit for notational brevity. Parameter identification is performed by minimizing the sum of the squared error between $\hat{u}_{\mathrm{L},k}$ and $u_{\mathrm{L},k}$ over a test dataset of $K$ samples. Our parameter estimation problem is given as

$$\hat{\boldsymbol{\psi}} \in \arg\min_{\boldsymbol{\psi}} \sum_{k=1}^{K} \left(u_{\mathrm{L},k} - \hat{u}_{\mathrm{L},k}(\boldsymbol{\psi})\right)^2, \tag{1.10}$$

which can be solved offline for experimental purposes [18] or in real-time as a recursive least squares problem [20–23]. The HPPC test is commonly used to generate the data required for this identification procedure, as it excites the battery with controlled current pulses at fixed SOC setpoints [16], enabling clear observation of both the OCV and the transient voltage dynamics associated with each RC pair. The test is carried out over several temperatures, and the resulting identified parameter functions $r_0(s, T)$, $r_j(s, T)$, and $c_j(s, T)$ are subsequently used to populate and dynamically update the system matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ in (1.8).

To carry out this parameter estimation problem, the HPPC test is used to provide adequate data. This test is a valuable technique in the battery industry and can be performed at either the single-cell or system level. It involves applying a series of constant current pulses to a battery at fixed SOC intervals and measuring its terminal voltage. The test can be repeated at desired temperatures, enabling clear observation of the change due to temperature on both the OCV and the transient voltage dynamics [16]. The resulting identified parameter functions $r_0(s,T)$, $r_j(s,T)$, and $c_j(s,T)$ are subsequently used to populate and dynamically update the system matrices **A**, **B**, **C**, and **D** in (1.8).

Several parameter estimation methods have been developed to resolve (1.10). They are grouped into the following broad categories: analytical methods, Kalman filter methods, optimization algorithms, least squares (LS), recursive least squares (RLS), and hybrid methods [18, 23]. A summary of the existing literature on the applicability, strengths, and limitations of each family of techniques is presented in Section 2.4.1.

## 1.2 Kalman Filtering for Nonlinear Systems

For battery SOC estimation, the standard KF is not well-suited to the nonlinear dynamics of LIBs [24]. In this section, we go over the theory of the EKF and UKF and present how they are modified to deal with the nonlinear dynamics.

### 1.2.1 Extended Kalman Filtering

The EKF operates similarly to the KF with a state prediction and a state correction step. However, the nonlinear dynamics are linearized at operating points of the state estimates. In the prediction step, the EKF forecasts the state and its uncertainty based on the nonlinear state transition model:

$$\hat{\mathbf{x}}_{k|k-1} = f(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{u}_k) + \mathbf{w}_k \tag{1.11}$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^\top + \mathbf{Q}_k, \tag{1.12}$$

where $\hat{\mathbf{x}}_{k|k-1}$ is the prior state estimate at time $k$ given measurements until $k-1$, $\mathbf{P}_{k|k-1} \in \mathbb{R}^{n_x \times n_x}$ is the corresponding prior error covariance matrix, and $\mathbf{Q}_k \in \mathbb{R}^{n_x \times n_x}$ represents the process noise covariance at time $k$. The nonlinear state transition is described by the mapping $f(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{u}_k)$ with $\mathbf{F}_k \in \mathbb{R}^{n_x \times n_x}$ denoting its Jacobian matrix evaluated at the posterior state estimate $\hat{\mathbf{x}}_{k-1|k-1}$. The predicted measurement vector, $\hat{\mathbf{y}}_k$ is given by:

$$\hat{\mathbf{y}}_k = h(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{v}_k, \tag{1.13}$$

where $h(\hat{\mathbf{x}}_{k|k-1})$ is the nonlinear measurement function, evaluated using the prior state estimate at time $k$. In the correction step, the EKF updates the Kalman gain, $\mathbf{K}_k \in \mathbb{R}^{n_x \times n_y}$, by using $\mathbf{H}_k \in \mathbb{R}^{n_y \times n_x}$, which is the Jacobian of $h(\cdot)$ at the prior state estimate, $\hat{\mathbf{x}}_{k|k-1}$. The state estimate is then updated using the residual error, defined as the difference between the observed measurement $\mathbf{y}_k$ and $\hat{\mathbf{y}}_k$. This leads to:

$$
\begin{aligned}
\mathbf{K}_k &= \mathbf{P}_{k|k-1}\mathbf{H}_k^\top \left(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^\top + \mathbf{R}_k\right)^{-1} \\
\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \left(\mathbf{y}_k - \hat{\mathbf{y}}_k\right) \\
\mathbf{P}_{k|k} &= \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{P}_{k|k-1},
\end{aligned}
$$

where $\hat{\mathbf{x}}_{k|k}$ and $\mathbf{P}_{k|k}$ are the posterior state estimate and error covariance matrix, respectively, and $\mathbf{R}_k \in \mathbb{R}^{n_y \times n_y}$ is the measurement noise covariance matrix. In this work, we only consider the measurement function $h(\hat{\mathbf{x}}_{k|k-1})$ to be nonlinear given the nature of the LIB load voltage dynamics given in (1.7). To obtain $\mathbf{H}_k$, we differentiate (1.7) with respect to the prior state estimate, which yields:

$$
\mathbf{H}_k = \begin{bmatrix} u'_{\mathrm{oc}}(s_k) & -1 & -1 \end{bmatrix}, \tag{1.14}
$$

where $u'_{\mathrm{oc}}(s_k)$ is the derivative of $u_{\mathrm{oc}}(s_k)$ with respect to $s_k$ evaluated at time $k$. Note that this linearization is loss-introducing and causes the EKF to no longer be an optimal estimator. Although the UKF often gives better results for nonlinear estimation problems [25], the use of the EKF allows us to extend the DRKF to nonlinear problems, as presented in Section 3.

### 1.2.2 Unscented Kalman Filtering

The UKF addresses the linearization limitations of the EKF by using a deterministic sampling approach known as the unscented transformation (UT) [26]. We offer a condensed explanation of its mechanism here to illustrate its differences with the EKF. Interested readers are referred to [26] for a full treatment of UKF.

Instead of approximating the nonlinear system with a first-order Taylor series, the UKF represents the state distribution as a Gaussian random variable (GRV) using a minimal set of carefully chosen sample points, referred to as sigma points, which completely capture the true mean and covariance of the GRV. When propagated through the true nonlinear system, these points capture the posterior mean and covariance accurately to the 3$^{\text{rd}}$ order for any nonlinearity [26]. The underlying mechanism of determining the posterior mean and covariance is the unscented transform [27]. Consider propagating $\mathbf{x}_k$ through a nonlinear measurement equation. Assume $\mathbf{x}_k$ has a mean $\bar{\mathbf{x}} \in \mathbb{R}^{n_x}$ and covariance $\mathbf{P}_{\mathbf{x}_k} \in \mathbb{R}^{n_x \times n_x}$, to calculate the statistics of $\mathbf{y}$ we formulate a matrix $\boldsymbol{\chi} \in \mathbb{R}^{n_x \times n_p}$, where $n_p = 2L + 1$ and

$L \in \mathbb{N}$ is a hyperparameter that determines the number of sigma vectors $\chi_i$ $i = 1, 2...., 2L$. The mean vector is placed in the first column, $\chi_0$, and the sigma vectors are varying shifts of $\bar{\mathbf{x}}$. Each sigma vector carries a corresponding weight, $W_i$ computed from $L$ and other hyperparameters [26]. The sigma points are then propagated through the measurement function:

$$\mathcal{Y}_i = h(\chi_i),\ i = 0, ....2L,$$

The mean and covariance for $\hat{\mathbf{y}}_k$ can then be approximated from the weights and the sigma points $\mathcal{Y}_{i,k}$. The Kalman gain is then computed directly using the two covariance matrices based on the sigma point propagations, and the state estimate can then be updated.

Note that no explicit calculation of Jacobian or Hessian matrices is necessary to implement this algorithm. Furthermore, the overall number of computations is the same order as the EKF [26]. In the context of LIB SOC estimation, for the 2-RC circuit, the Jacobian is not very costly and the EKF tends to be more computationally efficient [16, 18]. In Section 5, we discuss how this formulation drives the improved performance in estimation accuracy between the EKF and UKF, and present some ideas of how the DREKF could be improved utilizing this approach.

### 1.3 Distributionally Robust Kalman Filters

Many optimization problems in engineering, control, and energy systems rely on uncertain quantities such as demand, renewable generation, measurement noise, or system disturbances. Stochastic modelling approaches assume that the underlying probability distribution of the uncertainty is known exactly. In practice, however, this assumption is rarely justified. The true distribution is typically unknown and is estimated from limited or noisy data. As a result, stochastic optimization may yield solutions that perform well under the assumed distribution but degrade significantly when the model is misspecified. Robust Optimization addresses this issue by optimizing over the worst-case outcome for a particular problem, but it can be overly conservative if this worst-case is not met in reality. Distributionally robust optimization (DRO) provides a mathematical programming framework for decision-making under uncertainty that bridges the gap between stochastic and robust optimization [15]. Instead of assuming a known probability distribution, DRO defines an ambiguity set which contains a family of plausible distributions around the empirical data, and optimizes for the worst-case expectation within that set. This protects against distributional uncertainty while avoiding the over-conservatism of classic robust optimization [28, 29]. There are multiple ways to define the ambiguity set, namely:

- **Wasserstein balls:** Sets of distributions within a fixed Wasserstein distance from a nominal distribution;
- **$f$-divergence sets:** Based on statistical divergences such as KL divergence or $\chi^2$ distance;
- **Moments Based Similarity:** Considers all probability distributions that satisfy constraints on the moments (typically mean and covariance). There is no constraint on a specific distributional shape (e.g. Gaussian, uniform, etc.)

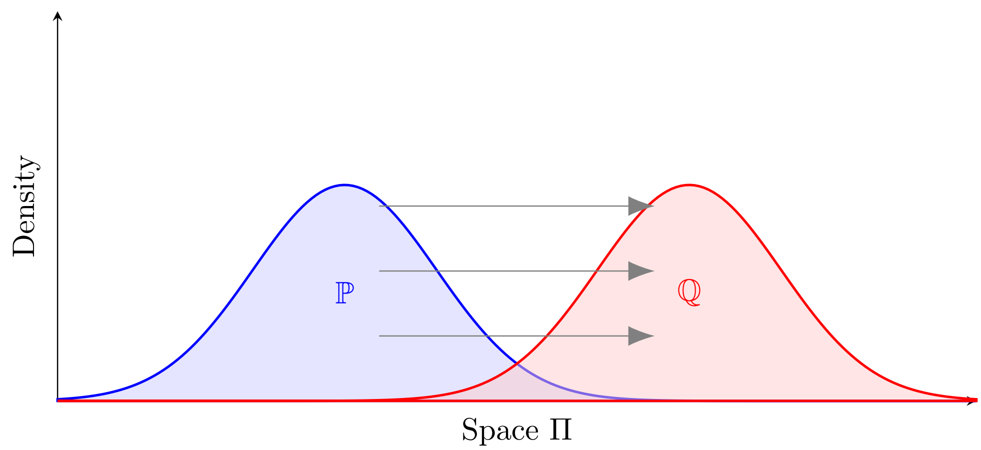



Figure 1.3 Illustration of the Wasserstein distance $\mathrm{W}_2(\mathbb{P}, \mathbb{Q})$ between the probability distributions $\mathbb{P}$ and $\mathbb{Q}$, represented as the minimal "mass transport" required to shape one distribution into the other.

Ambiguity sets based on the Wasserstein balls are particularly appealing, as they provide geometric interpretability. The Wasserstein distance measures the difference between probability distributions in terms of the minimal "transport cost" required to transform one distribution into another [29, 30], as illustrated in Figure 1.3. Furthermore, it renders computationally tractable formulations for problems involving quadratic objectives, such as our least mean-squared objective, which is presented in (3.1). Formally, the set of Borel probability measures with support $\Omega \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$ is denoted by $\mathcal{P}(\Omega)$, while the subset of measures with finite second-order moments is denoted by $\mathcal{P}_2(\Omega)$. For two probability distributions $\mathbb{P}, \mathbb{Q} \in \mathcal{P}(\Omega)$, the order-2 Wasserstein distance with respect to the Euclidean norm $||\cdot\|_2$, is defined as:

$$\mathrm{W}_2(\mathbb{P},\mathbb{Q}) = \inf_{\boldsymbol{\pi}\in\Pi(\mathbb{P},\mathbb{Q})} \left( \int_{\Omega\times\Omega} \|\mathbf{z}_1 - \mathbf{z}_2\|^2 \mathrm{d}\boldsymbol{\pi}(\mathbf{z}_1,\mathbf{z}_2) \right)^{\frac{1}{2}}, \tag{1.15}$$

where $\Pi(\mathbb{P},\mathbb{Q})$ denotes the set of all joint distributions on $\Omega\times\Omega$ with marginals $\mathbb{P}$ and $\mathbb{Q}$. Intuitively, the Wasserstein distance represents the minimum amount of "mass transport" required to morph one probability distribution into another. By constructing a Wasserstein ball of radius $\theta \geq 0$ centred around a nominal or assumed distribution $\hat{\mathbb{Q}}$, the distribution $\mathbb{P}$ is found within the ambiguity set, $\mathcal{B}_\theta(\hat{\mathbb{Q}})$. This is defined as

$$\mathcal{B}_\theta(\hat{\mathbb{Q}}) \coloneqq \left\{ \ \mathbb{P} \in \mathcal{P}(\Omega) | \ \mathrm{W}_2(\mathbb{P},\hat{\mathbb{Q}}) \leq \theta \ \right\}.$$

In the context of Kalman filtering, we define the empirical distribution $\hat{\mathbb{Q}} = \mathrm{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$ as a standard Gaussian distribution based on an assumed mean vector, $\boldsymbol{\mu}\in\mathbb{R}^{n_x}$ and covariance $\boldsymbol{\Sigma}\in\mathbb{R}^{n_x\times n_x}$. A typical assumption is to have zero-mean noise, while the covariance is determined by the standard deviation of the relevant measurement [16]. A larger $\theta$ implies greater protection against distributional shifts against this nominal distribution at the cost of solution conservatism, while a smaller $\theta$ yields solutions closer to the nominal mean, but may not be as robust [29]. Building on this construction of the ambiguity set, the DRO problem is given as:

$$\inf_{\mathbf{x}\in\chi} \sup_{\mathbb{P}\in\mathcal{B}_\theta(\hat{\mathbb{Q}})} \mathbb{E}^{\mathbb{P}}[g(\mathbf{x},\xi)]$$

where $\mathbf{x}\in\chi\subset\mathbb{R}^n$ denotes the decision vector constrained to the set $\chi$ and $g(\mathbf{x},\xi)$ is a cost function we seek to minimize, subject to a stochastic disturbance $\xi$. In Section 3.2, we demonstrate how this general DRO formulation is applied to the state-estimation problem. The problem can be viewed as a zero-sum game between a statistician choosing the estimator and a fictitious adversary choosing the distribution $\mathbb{P}$. The cost function is set as a minimum mean-square objective between the true state and the estimated state, and this yields an estimator suitable for the recursive nature of the KF [31, 32].

### 1.4 Contributions

This paper investigates the effectiveness of distributionally robust Kalman filtering for LIB SOC estimation under uncertain and potentially non-Gaussian noise. The main contributions are as follows:

1. We design a distributionally robust extended Kalman filter (DREKF) that integrates order-2 Wasserstein ambiguity sets into an EKF-based estimation framework, enabling improved robustness to noise distribution mismatch and measurement outliers in nonlinear settings.
2. We propose a computationally efficient DREKF for real-time LIB SOC estimation to hedge against non-Gaussian noise, in which worst-case covariance updates are selectively recomputed based on voltage residuals, preserving real-time feasibility despite the inclusion of semidefinite programs (SDPs).
3. We illustrate the performance of the DREKF on an open-source LIB dataset across multiple temperatures, initialization conditions, and corrupted measurement scenarios, demonstrating improved robustness relative to the EKF and UKF.

### 1.4.1 Publications

The main contributions presented in this Master's thesis have been submitted as a brief paper and is under review for publication in Transactions on Control Systems Technology:

- **Donald Fisher**, Keyhan Sheshyekani, Antoine Lesage-Landry, "Distributionally Robust Extended Kalman Filtering for State-of-Charge Estimation of Lithium-ion Batteries," Transactions on Control Systems Technology, April 2026

### 1.4.2 Presentations

This work is to be presented at the following conference:

- *Distributionally Robust Extended Kalman Filtering for State-of-Charge Estimation of Lithium-ion Batteries*, Optimization Days, May 2026, Montréal, Québec, Canada.

## 1.5 Thesis Outline

The remainder of the thesis is organized as follows. Chapter 2 presents a comprehensive literature review, covering the four main methods for LIB SOC estimation, as well as analyzing various parameter estimation methods. In Chapter 3, we introduce the theoretical foundations for the model and present the DREKF algorithm. Chapter 4 presents the results based on the open-source LG 18650HG2 Li-ion Battery dataset [33]. Finally, Chapter 5 summarizes the main contributions of the Master's thesis, discusses limitations, and outlines promising directions for future research.

# CHAPTER 2 LITERATURE REVIEW

This chapter surveys the literature surrounding all of the key components of LIB SOC estimation. From Section 2.1 to 2.3, we discuss work in three out of the four main categories of SOC estimation: CC, EMs, and data-driven methods. We establish the foundation for choosing to use ECMs as the basis of our model. In Section 2.4, we explore the literature to understand the trade-offs of various parameter estimation algorithms, and examine several filtering approaches. In Section 2.5, we delve into the existing work on distributionally robust state estimation, emphasizing its appeal for the battery SOC estimation problem. Finally, in Section 2.6 we summarize the main insights from the literature and present the research gap that is tackled in this thesis.

## 2.1 Coulomb Counting

CC is a simple approach to SOC estimation and places low requirements on the sensors and computation load [10]. However, the sensor accuracy and sampling interval introduce cumulative errors. Because it is an open-loop system with no feedback correction, wrong initial state values will lead to wrong estimated values, and the CC estimates tend to drift over time [10]. To overcome the error introduced by a wrong initial estimate, [34] proposes an improved CC method in which the capacity is periodically recalibrated to address any drift, updating the nominal capacity. While this improves on the traditional CC for aged batteries, it does not offer real-time correction, and so it is inferior to the other methods that will be discussed.

## 2.2 Electrochemical Models

EMs utilize internal electrochemical kinetics and Lithium-ion transport behaviour to predict the current SOC. The electrochemical model can characterize microscopic distributions of cell concentration, potential, current, and temperature and macroscopic characteristics such as cell voltage and current. Among many electrochemical models, the main EMs employed for LIBs are the pseudo-2-d (P2D) model [35] and the single particle model (SPM) [35–37].

Reference [38] first introduces the P2D model based on concentrated solution theory and porous electrode theory. It employs a set of coupled partial differential equations (PDEs) to track lithium-ion concentration and electrical potential within the battery. While the P2D model provides superior physical insight and accuracy compared to empirical models, its complexity serves as its primary limitation for real-time applications [35]. The high computational burden required to solve these nonlinear PDEs in real-time often exceeds the processing capabilities of a standard BMS [39].

Following, [40] proposes a reduced-order model of the P2D model known as the SPM. The SPM rests on two main assumptions: First, each electrode is modelled as two spherical particles in which intercalation and de-intercalation phenomena occur. Second, the variations in the electrolyte concentration and in the potential are ignored [35]. The SPM ignores the detailed distribution of local concentration and potential in the solution phase and instead accounts for a lumped solution resistance term [41]. The main advantage of this model is that it requires minimal computational effort, making it suitable for online estimation. However, the assumptions made in the SPM are only reasonable for low applied current densities, thin electrodes, and highly conductive electrodes. At high current densities, the simplification reduces the model's accuracy and does not capture the concentration dynamics as well as the P2D model [41,42], and it must be fine-tuned according to the electrolyte properties [35]. To tackle this, in [41] the SPM is extended to include thermal effects by adding an energy balance equation. These equations are coupled together by the nonlinear dependence of transport and kinetic parameters on cell temperature; so, no analytical solution can possibly be obtained for this equation system when the variation of the cell temperature is taken into account [41]. Therefore, although the accuracy at high C rates of this improved SPM compared to the standard SPM, the solution requires manually tuning for a termination point as it is solved for iteratively [41], which may not generalize well.

Overall, EMs generally mimic more precisely the behaviour of LIBs when compared to other approaches. However, these models are computationally more expensive and require numerous additional input parameters such as specific heat capacity, thermal diffusivity, and thermal conductivity. As such, this approach is typically only preferred in battery chemistries where the relationship between the SOC and OCV cannot be utilized, such as Lithium-sulphur [43] or in low C-rate use cases [35].

## 2.3 Data-Driven Methods

With the recent rise of machine learning (ML) and artificial intelligence in the last decade [44], data-driven approaches using traditional ML and deep learning (DL) algorithms for LIB SOC estimation have become increasingly popular in the literature. Many LIB experimental datasets have also been developed to support the research into data-driven algorithms. In this section, we provide an overview of the methods which have been implemented.

Traditional ML, DL, and transformers all utilize the known variables, such as voltage, current, and temperature, as well as manually crafted additional features to form a generalized function approximator. Given a sufficient amount of training data and an appropriate configuration, the SOC value can be predicted accurately without the need for a sophisticated physics-based model to explain the behaviour of the battery. Performance can also be improved through hyperparameter tuning, as well as by blending them with other optimization algorithms or even KFs.

Considering traditional ML techniques, [45] proposes a k-nearest neighbours method that uses charging and discharging data, decision tree algorithms and KFs to enable dynamic SOC estimation. The authors of [46] propose a well-generalized extreme learning machine, which improved on nonlinear fitting and real-time performance by utilizing kernel functions and optimization algorithms. Zhou et al. [47] blends a support vector machine (SVM) with a UKF, utilizing the SVM for parameter estimation and the UKF for SOC estimation, highlighting the versatility of ML methods.

Advancements in DL have led to its increased utilization for the LIB SOC estimation problem [48]. DL models such as multilayer perceptrons (MLPs), convolutional neural networks (CNNs), and recurrent neural networks (RNNs), excel in nonlinear fitting capabilities and can automatically learn feature representations, demonstrating strong performance in SOC estimation. MLPs map input variables and SOC by stacking fully connected layers [49]. RNNs are designed for sequential data processing and use recurrent connections to maintain a hidden state (memory) over time. This ability to understand temporal dependencies has made RNNs and their variants, including long short-term memory (LSTM) and gated recurrent units (GRUs), very effective for capturing the impact of historical measurements on the current SOC [50]. LSTMs and GRUs address the vanishing and exploding gradient problems in standard RNNs through gating mechanisms, making them particularly suited for processing long sequences in SOC estimation [51].

Transformer models have seen increased interest in their application to time series modelling problems. This is due to the success demonstrated in natural language processing and computer vision, attributed to the effectiveness of the multi-head self-attention mechanism [52]. In [53],the authors employ a transformer model to predict battery SOC using temperature, current, voltage and velocity. The transformer obtained impressive results, outperforming the benchmark LSTM models in the paper, and obtaining an RMSE of as low as 0.166% in one dataset.

Considering the blackbox nature of DL, some research has explored methods which aim to build on user trust. The authors in [54] combine a temporal convolutional network, a GRU, a deep neural network, and a dual attention mechanism to significantly improve the prediction accuracy of the remaining useful life of LIBs. Introducing the feature and temporal attention mechanisms further enhances the model's understanding of the battery performance degradation patterns, giving the model significant potential for application in battery health management and prediction [5]. However, transformers also suffer from large computational demands, and additional optimization strategies are needed for efficient BMS implementations [55]. Further development is also needed in order to ensure the transformer models can generalize well to more varied real-world datasets [55].

## 2.4 Equivalent Circuit Models

The application of ECMs reduces to two interconnected estimation subproblems: parameter estimation, which identifies the circuit element values governing the cell's electrical response, and the SOC estimation. Both subproblems share a common observational structure, relying on terminal voltage measurements to accurately identify their respective states. We now review the literature on each subproblem.

### 2.4.1 Parameter Identification

Parameter estimation plays a central role in battery modelling. In practice, this task is complicated by measurement noise and limited data availability, all of which can degrade estimation accuracy. As a result, a wide range of estimation techniques has been developed.

Analytical methods rely on equations to determine the values of each parameter. The Shepherd model is typically used in $LiFePO_4$ batteries [18, 56], but it is not well suited to other LIB chemistries and has difficulty determining the transient voltage behaviour of the LIB [57].

Kalman filter-based methods follow very closely from the formulation of the SOC estimation problem, and hold many of the same benefits and drawbacks that are discussed in this thesis. In our chosen 2-RC scenario, the model parameters $u_{\mathrm{oc}}(s), r_0, r_1, c_1, r_2, c_2$, form the state vector, while the battery current and voltage remain unchanged as the system input and output, respectively. The EKF can then be used again in this fashion to predict the optimal parameters. Dual filtering approaches have also been developed, whereby one Kalman filter predicts the optimal parameters and these are systematically integrated into another filter for the SOC state estimation. The work in [58] provides one such example, where a weighted EKF is introduced for the identification of parameters, along with a separate state-estimation algorithm. Overall, Kalman filters provide a highly adaptable form of real-time estimation, but are dependent on accurate initial models and noise assumptions, as is true for the SOC estimation problem.

Heuristics and meta-heuristic optimization methods offer both problem-specific and general approaches that have been applied to parameter estimation. Evolutionary algorithms such as genetic algorithms [59] or the coyote optimization algorithm [60] have recently been applied to the LIB parameter estimation problem. Overall, these methods are versatile in helping to find optimal parameter values and are user-friendly, making them an attractive choice for experimentation. They can, however, be quite computationally expensive for complex search spaces, especially as the dimension grows. They are also non-deterministic, have a certain black-box nature, and their convergence is not established [18].

The least-squares method is a classical and widely used estimation method. Based on (1.2), (1.3), it uses the Laplace transform to define an information vector. This vector is used to find the values that minimize the sum of the squared differences between the measured load voltage and the predicted voltage [18]. It is very popular for offline estimation; however, due to the requirement for batches of measurements to be available, it can only be used for offline estimation. The recursive least squares (RLS) represents the extension of this problem for real-time estimation and does so in a computationally efficient manner.References [20–22] all propose variants of the RLS specific for LIB ECMs, where a forgetting factor is introduced to combat the issue of data-saturation, which can occur due to the lengthy continuous nature of problems. These algorithms offer lightweight, accurate, and adaptable parameter estimation in real-time, and are very practical for real-time estimation, although they have some susceptibility to being unstable [18].

Finally, hybrid methods have also emerged recently. In [61] an approach merging the RLS and the KF is proposed to overcome the issues of data saturation in real-time. The RLS is used to identify a RC pair which corresponds to the charge transfer effect which takes place in under 10s, while an EKF tackles the long term diffusion based change. While this approach effectively improves identification accuracy, it significantly increases the number of OCV tests carried out at specific temperatures and aging stages [18], making data collection lengthier and more cumbersome.

Overall, parameter estimation is a very active research field, and each method provides certain strengths and weaknesses. While the RLS variants provide the most suitability for real-time estimation, we choose to use the optimization tools, and in particular MATLAB's `tfest` in this thesis. It allows for the specification of multiple numerical search methods, such as Levenberg-Marquardt or Gauss-Newton, to minimize the loss function [62], providing an accurate and convenient tool for rapid experimentation, and simplifying the research focus to the state estimation algorithm development. One caveat, however, is that as an iterative optimization tool, it can get stuck at local minima.

### 2.4.2 State estimation methods for SOC estimation

We now turn our attention to the literature on state estimation for SOC estimation. Model-based state estimation methods for LIBs are mainly filtering algorithms. As previously discussed, the KF is a highly adopted method for SOC estimation due to its recursive nature, its ability to handle sensor noise, and to self-correct. However, challenges arise due to the nonlinear nature of LIBs and the assumption of a known Gaussian noise process. Several improved filters have been proposed to tackle this. These include, but are not limited to, the EKF, the UKF, the $H_\infty$ filter, and the particle filter (PF). We now survey their advantages and limitations in the literature.

#### Extended Kalman Filter

Plett et al. [63] first introduced the EKF for LIB SOC estimation, and the EKF has been widely adopted for battery SOC estimation since. However, in the case of a highly nonlinear system, the linearization around every predicted system state may fail to represent the corresponding nonlinear system accurately and provide insufficient robustness. Reference [64] proposes incorporating variable parameters within the EKF framework, achieving a reduction in maximum estimation error, particularly at high C-rates. Then [65] proposes an adaptive iterative EKF, which introduces an adaptive maximum correlation entropy criterion to replace the minimum mean square error criterion in the traditional EKF to enhance the robustness to abnormal data and non-Gaussian noise [5]. Reference [66] introduces a multi-modal EKF, where the temperature and current effects are separated into different models, displaying improved results at extreme temperatures. However, this work focuses on Lithium titanate batteries and is not explicitly generalizable to other Li-ion chemistry's [5, 66].

#### Unscented Kalman Filters

Due to its use of the Unscented Transform [27], the UKF operates at a slightly higher computational cost than the EKF; however, because it avoids the Jacobian approximation, it avoids the linearization error and tends to outperform the EKF [5]. To strengthen its use case with respect to fixed noise assumptions, adaptive UKF (AUKF) variants can be implemented. Adaptive KFs tune the process noise covariance $\mathbf{Q}_k$, the measurement noise covariance $\mathbf{R}_k$, and the error covariance $\mathbf{P}_k$ in real-time through the addition of adaptive-law covariance matching [25], which is typically done through innovation-based sequences or residual based sequences [67]. Reference [25] introduces an AUKF that updates the system and measurement noise covariances in real time, yielding slightly improved performance over standard EKF and UKF implementations. However, AUKF's SOC estimation accuracy and stability can degrade given the pseudo-positive definiteness of the covariance matrix and inaccurate noise statistics [68]. To overcome this, [69] develops a new exponential adaptive strong tracking UKF to address issues of slow tracking speed by utilizing an evanescant factor, and altering the noise covariance through a Sage-Husa algorithm [69].

The UKF suffers from stability issues occasionally. This is attributed to an ill-conditioned covariance matrix in the iterative process, because the Cholesky decomposition requires that the covariance matrix be positive definite [70]. To address the covariance instability, the square root UKF (SRUKF) is proposed, as it does not require the reconstruction of the covariance matrix and, therefore, the stability issue is mitigated. The author of [68] proposes an adaptive SRUKF and a dual-coefficient tracker, which demonstrates improved accuracy greater than 3% compared to the UKF and SRUKF, although the dual-coefficient tracker showed some degradation and was not as accurate in updating the noise covariance. Reference [71] implements an adaptive fractional-order SRUKF, using a fractional-order model to represent the ECM, and demonstrates improved robustness and accuracy to heavy-tailed measurement noise. While these methods are encouraging, the SRUKF does not completely resolve the problem of pseudo-positive definiteness due to its nonlinear state calculation [68], and so stability is not generally guaranteed.

In [72], a fuzzy robust two-stage unscented Kalman filter (FRTSUKF) is implemented to improve the SOC estimation accuracy, whereby the model uncertainties are accounted for without the need for their statistical characteristics. The algorithm uses a fuzzy system to recursively adjust the covariance matrix of the measurement noise, making the estimation process more accurate and practical. In their results, they demonstrate a improved accuracy by as much as 1% compared to a robust UKF baseline. However, the design and parameter selection of fuzzy logic systems require optimization based on empirical or trial-and-error methods, which limits their general application without sufficient expertise [5].

In essence, AUKFs offer improved performance when compared to the standard EKF and UKF [68]. They can maintain accuracy as the battery ages, the noise characteristics change, and the thermal environment changes reducing the need for manual tuning of noise parameters. However, they can diverge if the adaptation logic reacts too aggressively to transient outliers or sensor faults, in which case SRUKF or FRTSUKF may be more suitable [71, 72]. Critically, all aforementioned filters still assume that the measurement and process noise follow a Gaussian distribution, limiting their applicability in deployments where this may not be the case.

### 2.4.3 Robust Filtering

Robust filters provide an appealing alternative to the limitations of Kalman filters as they provide an ability to account for the model error, uncertainty and non-Gaussian noise. The work of [73] first proposes the $H_\infty$ filter, which solves the uncertainty issue by using a mini-max cost function, minimizing the estimation error for the worst-case noise scenario without needing to know the noise distribution. However, similar to the KF, the nonlinear relationship between OCV and SOC renders the filter unsuitable for this problem. To this extent, [74] proposes a $H_\infty$ which utilizes radial basis functions as a way to approximate $u_{\mathrm{oc}}(s)$, demonstrating that the $H_\infty$ filter could be modified for the nonlinear SOC estimation problem. Later, [75] proposes a strong tracking approach to the $H_\infty$ filter by introducing a suboptimal fading factor to track some of the methods' more rapidly changing parameters. This approach demonstrates an overall improved robustness over the EKF benchmark, and with comparable computational cost, although it also demonstrates some instability at 0°C.

PFs are also an effective alternative to KFs. They are effective for nonlinear systems, and can cope with non-Gaussian noise distributions [76]. The main idea of the PF is to utilize a sequence of scaled, randomly selected particles to estimate the probability density function on any state space model. The position and state of the particles are then adjusted depending on the observed values to reach optimal variation estimates from the sample mean. The author of [77] applies the PF to a SOC estimation problem for supercapacitors, utilizing a fractional order model, where a common phase angle element is used instead of a pure capacitive element. The authors of [78] design a dual estimation PF based on a second-order ECM to estimate both the SOC and state-of-health (SOH), obtaining accurate MAE and RMSE values across temperatures and testing datasets. However, accurate PFs require a large number of particles to be generated. The consequence of this is that the computational cost is very high, and so KFs are typically preferred for EV BMSs. Additionally, the PFs require resampling strategies to counter the effect of sample degradation, and there is also room for improvement in this area when it comes to computational efficiency [79].

## 2.5 Distributionally Robust Kalman Filtering

In recent years, distributionally robust state estimation (DRSE) has emerged as a powerful framework for handling ambiguities in the noise distributions. Rather than relying on exact knowledge of the process and measurement noise distributions, we instead seek state estimators that perform well under the worst-case distributions within a pre-specified ambiguity set centred around a nominal model. Because DRSE looks at the worst-case distribution instead of the worst-case value, it offers the possibility of having tighter bounds compared to $H_\infty$ filters while still being robust to non-Gaussian and uncertain noise distributions. Within the literature, [31] established the foundational framework for a DRKF in the linear setting. Building upon this, [80] developed formulations for both finite- and infinite-horizon problems, while [32] proposed a steady state DRKF, whereby theoretical conditions under which the Kalman gain may be computed offline, allowing for the elimination of continuously solving a semidefinite program (SDP) at each timestep. However, these formulations have only been introduced for linear state estimation problems.

When considering nonlinear problems, [81] proposes a distributionally robust PF, whereby the worst-case likelihood distribution of the prior state particles is found from a ball centred around a nominal likelihood distribution, and the size of this ball is defined by the Wasserstein distance. However, considering that the computational cost of PFs is higher than that of KFs, we focus on the latter to ensure our approach's compatibility with real-time SOC estimation. To the best of our knowledge at the time of writing, no prior work exists where the DRKF has been modified to be applicable in the nonlinear setting.

### 2.6 Literature Summary

Due to the widespread adoption of LIBs, substantial research has been devoted to the development of SOC estimation techniques. Among the available techniques, CC, direct OCV measurement and EMs are typically ill-suited to the requirements of real-time estimation for LIBs. Although ML and DL methods have been shown to achieve impressive results with relatively straightforward implementations using easily measurable data, these methods often depend on extensive feature engineering processes and exhibit limited generalization capability in complex scenarios [82]. Another issue is the amount of data needed for training the models, and the black-box nature of data-driven methods means that these models are difficult to interpret, and their reliability cannot be guaranteed. Additionally, substantial computational cost and memory are required to train models, and this can pose difficulties when deploying such models on resource-constrained devices, such as the embedded systems used for battery management systems. Combined, these considerations currently limit data-driven techniques from deployment in real-time SOC estimation, but it is considered a field with high potential, and industry players are currently investing in hybrid Kalman filtering and ML strategies as ML techniques continue to improve [83, 84]. Future work requires further efforts to optimize the model performance, reduce the complexity, and improve the generalization ability of data-driven methods.

As a result, KFs based on ECMs are the most commonly employed models in practice [9, 85]. This preference is attributable to their relative simplicity, moderate computational requirements, and reduced dependence on extensive training data, while still capturing the dominant battery dynamics necessary for real-time estimation. They also offer more explainable results in comparison to the blackbox data-driven methods.

ECM SOC estimation techniques are largely built upon Kalman filtering frameworks, with the most popular being the EKF and the UKF. Adaptive variants of these filters have shown improved results over the standard EKF and UKF when considering shifts in the noise distributions from the nominal assumptions. However, a fundamental limitation shared by most of these approaches is the assumption of a known, Gaussian process and measurement noise. In practice, sensor errors, modelling inaccuracies, aging effects, and external disturbances often lead to non-Gaussian and potentially heavy-tailed noise distributions, especially in EVs. Robust filters are a viable tool for SOC estimation and provide an effective way to handle model uncertainty and non-Gaussian noise assumptions. On the one hand, the $H_\infty$ filter is more robust to uncertainties and model errors than the Kalman filter, which can be more sensitive to such errors. On the other hand, the $H_\infty$ filter is more complex to model and implement for any system. PFs relax the Gaussian assumption and provide a fully Bayesian solution, offering a robust alternative even in the nonlinear setting. They are, however, typically computationally intensive and may be impractical for a real-time BMS with limited processing capability.

In this work, we address the limitations associated with the fixed, Gaussian noise assumption of conventional KFs, which can lead to larger estimation errors if the true noise distribution does not match the nominal assumption, and can occasionally cause filter divergence. We choose to further develop the DRKF framework to avoid the additional computational burden of PFs and the overly conservative nature of the $H_\infty$ filter. We combine the DRKF and EKF to form our new DREKF, which extends the DRKF to the nonlinear setting. We subsequently apply the DREKF to the LIB SOC estimation problem. This framework enables improved robustness to model mismatch and non-Gaussian disturbances while maintaining computational tractability suitable for real-time implementation. This framework offers flexibility when compared to the $H_\infty$ filter, as the degree of robustness of the filter can be tuned by adjusting the size of the ambiguity set used through the tuning of the $\mathrm{W}_2$ distance.

# CHAPTER 3 PROPOSED METHODOLOGY

We now formulate the proposed DREKF.

## 3.1 Notation

Let $\mathcal{S}^n$ denote the set of all symmetric matrices in $\mathbb{R}^{n\times n}$. We use $\mathcal{S}^n_+ \subset \mathcal{S}^n$ and $\mathcal{S}^n_{++} \subset \mathcal{S}^n$ to represent the cone of symmetric positive semidefinite and positive definite matrices in $\mathcal{S}^n$, respectively. For any $\mathbf{A}, \mathbf{B} \in \mathcal{S}^n$, $\mathbf{A} \succeq \mathbf{B}$ and $\mathbf{A} \succ \mathbf{B}$ mean that $\mathbf{A} - \mathbf{B} \in \mathcal{S}^n_+$ and $\mathbf{A} - \mathbf{B} \in \mathcal{S}^n_{++}$, respectively. The set of Borel probability measures with support $\Omega \subseteq \mathbb{R}^n$ is denoted by $\mathcal{P}(\Omega)$, while the subset of measures with finite second-order moments is denoted by $\mathcal{P}_2(\Omega)$.

## 3.2 Problem Setup

Consider the state-space problem defined similarly to (1.8). We also assume that the nominal distributions for the initial state $\hat{\mathbf{x}}_0$, process noise $\hat{\mathbf{w}}_k$, and measurement noise $\hat{\mathbf{v}}_k$ for all $k$ are Gaussian with distributions $\hat{\mathbb{Q}}_{\mathbf{x}_0} = \mathrm{N}(\hat{\mathbf{x}}_0, \hat{\boldsymbol{\Sigma}}^-_{\mathbf{x}_0})$, $\hat{\mathbb{Q}}_{\mathbf{w}_k} = \mathrm{N}(\hat{\mathbf{w}}_k, \hat{\boldsymbol{\Sigma}}_{\mathbf{w}_k})$, and $\hat{\mathbb{Q}}_{\mathbf{v}_k} = \mathrm{N}(\hat{\mathbf{v}}_k, \hat{\boldsymbol{\Sigma}}_{\mathbf{v}_k})$, respectively, where $\hat{\mathbf{x}}_0, \hat{\mathbf{w}}_k \in \mathbb{R}^{n_x}$ and $\hat{\mathbf{v}}_k \in \mathbb{R}^{n_y}$ are the mean vectors, while $\hat{\boldsymbol{\Sigma}}^-_{\mathbf{x}_0}$, $\hat{\boldsymbol{\Sigma}}_{\mathbf{w}_k} \in \mathcal{S}^{n_x}_+$ and $\hat{\boldsymbol{\Sigma}}_{\mathbf{v}_k} \in \mathcal{S}^{n_y}_+$ are the nominal covariance matrices. In practice, these nominal distributions are estimates as the true distributions governing $\mathbf{x}_0$, $\mathbf{w}_k$, and $\mathbf{v}_k$ are not directly observable. This motivates the use of an estimator that performs well if the true distributions deviate from the nominal assumptions.

Because the state estimate and measurement are coupled through the observation model, a worst-case distribution is sought over their joint distribution rather than over each marginal independently. At each time step $k$, the prediction step of the filter propagates both the state estimation and process noise uncertainty through the state transition dynamics, such that the process noise $\mathbf{w}_k$ is incorporated into a prior state estimate distribution, $\mathbb{P}_{\hat{\mathbf{x}}_{k|k-1}} \in \mathcal{P}_2(\mathbb{R}^{n_x})$ [31, 32]. The measurement noise distribution, defined as $\mathbb{P}_{\mathbf{v}_k} \in \mathcal{P}_2(\mathbb{R}^{n_y})$, remains independent [31, 32]. We write the joint distribution as $\mathbb{P}_k = \mathbb{P}_{\hat{\mathbf{x}}_{k|k-1}} \otimes \mathbb{P}_{\mathbf{v}_k}$, where $\otimes$ denotes the product measure. Similarly, we define the nominal joint distribution as $\hat{\mathbb{Q}}_k = \hat{\mathbb{Q}}_{\hat{\mathbf{x}}_{k|k-1}} \otimes \hat{\mathbb{Q}}_{\mathbf{v}_k}$. To account for the uncertainty between $\mathbb{P}_k$ and $\hat{\mathbb{Q}}_k$, we consider an ambiguity set $\mathcal{B}_{\boldsymbol{\theta}}(\hat{\mathbb{Q}}_k) \subset \mathcal{P}_2(\mathbb{R}^{n_x}) \otimes \mathcal{P}_2(\mathbb{R}^{n_y})$ consisting of all joint probability distributions that are sufficiently close to the nominal distribution, $\hat{\mathbb{Q}}_k$. We control this distance through the use of the order-2 Wasserstein distance with respect to the Euclidean norm $||\cdot||_2$ [86], denoted as $\mathrm{W}_2(\mathbb{P}_1, \mathbb{P}_2)$ for two probability measures $\mathbb{P}_1$ and $\mathbb{P}_2$. Formally, $\mathcal{B}_{\boldsymbol{\theta}}(\hat{\mathbb{Q}}_k)$ is defined as:

$$\mathcal{B}_{\boldsymbol{\theta}}(\hat{\mathbb{Q}}_k) := \Big\{\mathbb{P}_k = \mathbb{P}_{\hat{\mathbf{x}}_{k|k-1}} \otimes \mathbb{P}_{\mathbf{v}_k} \,|\, \mathbb{P}_{\hat{\mathbf{x}}_{k|k-1}} \in \mathcal{P}_2(\mathbb{R}^{n_x}),\ \mathbb{P}_{\mathbf{v}_k} \in \mathcal{P}_2(\mathbb{R}^{n_y}),$$
$$\mathrm{W}_2(\mathbb{P}_{\hat{\mathbf{x}}_{k|k-1}}, \hat{\mathbb{Q}}_{\mathbf{x}_{k|k-1}}) \leq \theta_{\mathbf{x}},\ \mathrm{W}_2(\mathbb{P}_{\mathbf{v}_k}, \hat{\mathbb{Q}}_{\mathbf{v}_k}) \leq \theta_{\mathbf{v}}\Big\},$$

where $\boldsymbol{\theta} = (\theta_{\mathbf{x}}, \theta_{\mathbf{v}})^\top \geq \mathbf{0}$ are the radii of the Wasserstein ball for the process and measurement noise, respectively. With the ambiguity set defined, we seek a state estimator that minimizes the worst-case mean-square error across all distributions with respect to the ambiguity set. We define this distributionally robust minimum mean-square problem as:

$$\min_{\phi_k \in \mathcal{F}_k} \max_{\mathbb{P}_k \in \mathcal{B}_\theta(\hat{\mathbb{Q}}_k)} \mathbb{E}^{\mathbb{P}_k}\left[\|x_k - \phi_k(y_k)\|^2\right] \tag{3.1}$$

where $\phi_k$ denotes the state estimator, $\mathcal{F}_k$ denotes the set of all admissible estimators, and $\mathbb{Q}_k$ represents the joint distribution of the state and measurement noise. The inner maximization selects the least-favourable distribution within the ambiguity set $\mathcal{B}_\theta(\hat{\mathbb{Q}}_k)$, and the radii $\theta_{\mathbf{x}}$ and $\theta_{\mathbf{v}}$ act as hyperparameters that govern the size of the ambiguity set. As $\theta_{\mathbf{x}}, \theta_{\mathbf{v}} \to 0$, the ambiguity set collapses to the nominal distribution and the problem recovers the standard minimum mean-square error solution, while larger radii yield a more robust but potentially overly conservative estimator [31]. Problem (3.1) is generally nonconvex. However, [32, 86] show that if the nominal distributions $\hat{\mathbb{Q}}_{\mathbf{x}_0}$, $\hat{\mathbb{Q}}_{\mathbf{w}_k}$, and $\hat{\mathbb{Q}}_{\mathbf{v}_k}$ are assumed to be Gaussian, then (3.1) has a tractable reformulation. The resulting estimator preserves the recursive structure of the KF, but replaces the nominal covariances for the process and measurement noise with worst–case covariances, in line with the robustness consideration.

We now present the DRKF [32, 87] and then devise the DREKF in Section 3.3. The DRKF begins with a prediction step: At time $k$ and given the previous posterior estimate $(\hat{\mathbf{x}}_{k-1|k-1}, \hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k-1|k-1}})$, the prior state estimate, $\hat{\mathbf{x}}_{k|k-1}$, and the pseudo-nominal prior state covariance, $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k|k-1}}$ are calculated as:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \tag{3.2}$$
$$\hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k|k-1}} = \mathbf{A}\hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k-1|k-1}}\mathbf{A}^\top + \hat{\boldsymbol{\Sigma}}_{\mathbf{w}_k}. \tag{3.3}$$

Then, instead of directly using the pseudo-nominal prior state covariance $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k|k-1}}$ and measurement covariance $\hat{\boldsymbol{\Sigma}}_{\mathbf{v}_k}$, the DR formulation computes the least–favourable covariances by solving the following SDP [32]:

$$\max_{\substack{\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}},\, \boldsymbol{\Sigma}_{\mathbf{x}_{k|k}},\, \boldsymbol{\Sigma}_{\mathbf{v}_k} \\ \mathbf{Y},\, \mathbf{Z}}} \quad \mathrm{Tr}[\boldsymbol{\Sigma}_{\mathbf{x}_{k|k}}] \tag{3.4a}$$

$$\text{s.t.} \quad \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}} - \boldsymbol{\Sigma}_{\mathbf{x}_{k|k}} & \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}} \mathbf{C}_k^\top \\ \mathbf{C}_k \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}} & \mathbf{C}_k \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}} \mathbf{C}_k^\top + \boldsymbol{\Sigma}_{\mathbf{v}_k} \end{bmatrix} \succeq 0 \tag{3.4b}$$

$$\begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k|k-1}} & \mathbf{Y} \\ \mathbf{Y}^\top & \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}} \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{\mathbf{v}_k} & \mathbf{Z} \\ \mathbf{Z}^\top & \boldsymbol{\Sigma}_{\mathbf{v}_k} \end{bmatrix} \succeq 0 \tag{3.4c}$$

$$\mathrm{Tr}[\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}} + \hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k|k-1}} - 2\mathbf{Y}] \leq \theta_{\mathbf{x}}^2 \tag{3.4d}$$

$$\mathrm{Tr}[\boldsymbol{\Sigma}_{\mathbf{v}_k} + \hat{\boldsymbol{\Sigma}}_{\mathbf{v}_k} - 2\mathbf{Z}] \leq \theta_{\mathbf{v}}^2, \tag{3.4e}$$

$$\boldsymbol{\Sigma}_{\mathbf{x}_{k|k}},\ \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}},\ \hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k|k-1}} \in \mathcal{S}_+^{n_x},\ \boldsymbol{\Sigma}_{\mathbf{v}_k} \in \mathcal{S}_{++}^{n_y}, \tag{3.4f}$$

$$\mathbf{Y} \in \mathbb{R}^{n_x \times n_x},\ \mathbf{Z} \in \mathbb{R}^{n_y \times n_y} \tag{3.4g}$$

where $\boldsymbol{\Sigma}_{\mathbf{x}_{k|k}}, \boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}, \hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k|k-1}} \in \mathcal{S}_+^{n_x},\ \boldsymbol{\Sigma}_{\mathbf{v}_k} \in \mathcal{S}_{++}^{n_y},\ \mathbf{Y} \in \mathbb{R}^{n_x \times n_x},\ \mathbf{Z} \in \mathbb{R}^{n_y \times n_y}$. The decision variables $\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}$ and $\boldsymbol{\Sigma}_{\mathbf{x}_{k|k}}$ represent the prior and posterior state covariances, respectively, while $\boldsymbol{\Sigma}_{\mathbf{v}_k}$ represents the measurement covariance. The objective (3.4a) finds the pair of $\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}$ and $\boldsymbol{\Sigma}_{\mathbf{v}_k}$ that maximizes $\boldsymbol{\Sigma}_{\mathbf{x}_{k|k}}$ within the ambiguity set. Constraint (3.4b) ensures that $\boldsymbol{\Sigma}_{\mathbf{x}_{k|k}}$ respects the Kalman filter mechanism. The auxiliary variables $\mathbf{Y}$ and $\mathbf{Z}$ and the resulting constraints (3.4c)−(3.4e) follow from the use of Schur complement arguments and Gelbrich bounds to convert the original infinite-dimensional optimization problem into a convex finite-dimensional problem [88], and ensure that $\boldsymbol{\Sigma}_{\mathbf{x}_{k|k-1}}$ and $\boldsymbol{\Sigma}_{\mathbf{v}_k}$ lie within the specified Wasserstein radii of the ambiguity set $\mathcal{B}_{\boldsymbol{\theta}}(\hat{\mathbb{Q}}_k)$. The reader is referred to [88] for the full derivation of (3.4). Solving (3.4) yields the least-favourable covariances, denoted as $\boldsymbol{\Sigma}^*_{\mathbf{x}_{k|k-1}}$ and $\boldsymbol{\Sigma}^*_{\mathbf{v}_k}$ which are used to determine the robust Kalman gain as

$$\mathbf{K}_k^* = \boldsymbol{\Sigma}^*_{\mathbf{x}_{k|k-1}} \mathbf{C}_k^\top \Big( \mathbf{C}_k \boldsymbol{\Sigma}^*_{\mathbf{x}_{k|k-1}} \mathbf{C}_k^\top + \boldsymbol{\Sigma}^*_{\mathbf{v}_k} \Big)^{-1}. \tag{3.5}$$

The posterior state estimate and covariance are determined via

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k^* \Big( \mathbf{y}_k - \hat{\mathbf{y}}_k \Big) \tag{3.6}$$

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k|k}} = (\mathbf{I} - \mathbf{K}_k^* \mathbf{C}_k) \boldsymbol{\Sigma}^*_{\mathbf{x}_{k|k-1}}. \tag{3.7}$$

where $\mathbf{I} \in \mathbb{R}^{n_x \times n_x}$ is the identity matrix. The Wasserstein radii $\theta_{\mathbf{x}}$ and $\theta_{\mathbf{v}}$ determine the estimator's conservativeness and serve as tunable hyperparameters selected from the data. Notably, for quadratic loss functions under a $\mathrm{W}_2$-ambiguity set centred at a Gaussian distribution, the worst-case distribution is known to remain Gaussian. This follows from the fact that the optimal adversarial distribution is an affine mapping of the nominal one, as demonstrated in [89].

### 3.3 Distributionally Robust Extended Kalman Filter

The DRKF framework of Section 3.2 assumes a linear output mapping $\mathbf{C}_k$. To extend DRKF to the nonlinear setting, we adopt a methodology akin to EKF, which we refer to as DREKF. We begin by assuming that $\mathbf{x}_k$ and $\mathbf{y}_k$ are governed by the nonlinear state and measurement functions $f(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{u}_k)$ and $h(\hat{\mathbf{x}}_{k|k-1})$, respectively. The prior state estimate, $\hat{\mathbf{x}}_{k|k-1}$, is computed as in (1.11). The nonlinear state transition matrix is obtained via the Jacobian $\mathbf{F}_k = \frac{\partial f}{\partial \hat{\mathbf{x}}}\Big|_{\hat{\mathbf{x}}_{k-1|k-1}}$. The pseudo-nominal prior state covariance $\hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k|k-1}}$ is obtained by using $\mathbf{F}_k$ to approximate the linear dynamics. The prediction step becomes:

$$\hat{\mathbf{x}}_{k|k-1} = f(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{u}_k) + \mathbf{w}_k \tag{3.8}$$

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k|k-1}} = \mathbf{F}_k \boldsymbol{\Sigma}_{\mathbf{x}_{k-1|k-1}} \mathbf{F}_k^\top + \hat{\boldsymbol{\Sigma}}_{\mathbf{w}_k}. \tag{3.9}$$

Similarly, the linearized time-varying measurement matrix is obtained via the Jacobian $\mathbf{H}_k = \frac{\partial h}{\partial \hat{\mathbf{x}}}\Big|_{\hat{\mathbf{x}}_{k|k-1}}$. Matrix $\mathbf{H}_k$ replaces matrix $\mathbf{C}_k$ within the SDP constraints (3.4b), the robust gain calculation (3.5), and the posterior state covariance update (3.7). Because $\mathbf{H}_k$ is linear, it preserves the Kalman filter's update mechanism for propagating the uncertainty in constraint (3.4b), as well as when updating the Kalman gain and the posterior state covariance [87]. The updated robust Kalman gain is then computed as:

$$\mathbf{K}_k^* = \boldsymbol{\Sigma}^*_{\mathbf{x}_{k|k-1}} \mathbf{H}_k^\top \Big(\mathbf{H}_k \boldsymbol{\Sigma}^*_{\mathbf{x}_{k|k-1}} \mathbf{H}_k^\top + \boldsymbol{\Sigma}^*_{\mathbf{v}_k}\Big)^{-1}, \tag{3.10}$$

and the posterior state and covariance updates are now computed using the linearized approximation:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k^* \Big(\mathbf{y}_k - \hat{\mathbf{y}}_k\Big) \tag{3.11}$$

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{x}_{k|k}} = (\mathbf{I} - \mathbf{K}_k^* \mathbf{H}_k) \boldsymbol{\Sigma}^*_{\mathbf{x}_{k|k-1}}. \tag{3.12}$$

In the context of LIB SOC estimation, nonlinearity is driven solely by the OCV polynomial $u_{\mathrm{oc}}(s)$. Because the state-transition function is linear, $\mathbf{F}_k = \mathbf{A} \ \forall \ k$ while $\mathbf{H}_k$ is defined by (1.14).

The DREKF is presented in Algorithm 1. We begin by predicting the state prior estimate, $\hat{\mathbf{x}}_{k|k-1}$, and the pseudo-nominal prior state covariance matrix, $\hat{\mathbf{\Sigma}}_{\mathbf{x}_{k|k-1}}$ via (3.8) and (3.9) (Line 7). Matrix $\mathbf{C}_k$ is approximated by $\mathbf{H}_k$ (Line 9) and we substitute it into (3.4b) to solve the SDP (3.4) (Lines 11−12). The robust Kalman gain is then computed via (3.10) and the state estimate and covariance are updated via (3.11) and (3.12) (Line 17).

**Algorithm 1** Distributionally Robust Extended Kalman Filter (DREKF)

**Inputs:**
1: Wasserstein radii $\theta_{\mathbf{x}}, \theta_{\mathbf{v}}$.
2: Covariances $\hat{\mathbf{\Sigma}}_{\mathbf{w}}, \hat{\mathbf{\Sigma}}_{\mathbf{v}}.\hat{\mathbf{\Sigma}}_{\mathbf{x}_0}$.
3: Initial state estimate and noise means $\hat{\mathbf{x}}_0, \hat{\mathbf{w}}_k, \hat{\mathbf{v}}_k$.
4: Update Tolerance $\epsilon$
**Output:** State estimates $\{\hat{\mathbf{x}}_k\}_{k=1}^{T}$
5: **for** $k = 1$ to $T$ **do**
6: Compute Jacobian $\mathbf{F}_k = \frac{\partial f}{\partial \hat{\mathbf{x}}}\Big|_{\hat{\mathbf{x}}_{k-1|k-1}}$.
7: Predict $\hat{\mathbf{x}}_{k|k-1}$ and $\hat{\mathbf{\Sigma}}_{\mathbf{x}_{k|k-1}}$ via (3.8) and (3.9).
8: Obtain measurement $\mathbf{y}_k$ and predict $\hat{\mathbf{y}}_k$ via (1.7).
9: Compute $\mathbf{H}_k = \frac{\partial h}{\partial \hat{\mathbf{x}}}\Big|_{\hat{\mathbf{x}}_{k|k-1}}$.
10: **if** $\|\mathbf{y}_k - \hat{\mathbf{y}}_k\| \geq \epsilon$ or $k = 1$ **then**
11: Substitute $\mathbf{H}_k$ into (3.4b).
12: Solve SDP (3.4) to obtain $\mathbf{\Sigma}^*_{\mathbf{x}_{k|k-1}}, \mathbf{\Sigma}^*_{\mathbf{v}_k}$.
13: **else**
14: $\mathbf{\Sigma}^*_{\mathbf{x}_{k|k-1}}, \mathbf{\Sigma}^*_{\mathbf{v}_k} = \mathbf{\Sigma}^*_{\mathbf{x}_{k-1|k-2}}, \mathbf{\Sigma}^*_{\mathbf{v}_{k-1}}$.
15: **end if**
16: Compute $\mathbf{K}^*_k$ via (3.10).
17: Compute $\hat{\mathbf{x}}_{k|k}, \hat{\mathbf{\Sigma}}_{\mathbf{x}_{k|k}}$ via (3.11) and (3.12).
18: **end for**

A drawback of the DREKF is that it requires $\mathbf{H}_k$ to be updated at each time step $k$. This in turn necessitates recomputing the worst-case covariance at each time step, and precludes the use of a more efficient offline (3.4), where the Kalman gain is computed once [32]. To improve computational efficiency, we draw inspiration from works based on event-triggered state estimation [90, 91] and assume that if the predicted measurement is very close to the actual measurement, the worst-case distribution from the previous timestep can be considered a reasonable approximation, and we do not require a recomputation of (3.4). Therefore, we recompute the worst-case covariances only when the difference between the measured terminal voltage and the estimated value exceeds a certain tolerance, $\epsilon \geq 0$ (Lines 10 and 13). We highlight this modification in grey in Algorithm 1. This additional hyperparameter can be tuned, depending on the desired accuracy and computational efficiency gain of the user and its effect is studied in Section 4.4. Setting $\epsilon = 0$, effectively ensures that (3.4) is recomputed at every timestep, leading to the nominal DREKF.

# CHAPTER 4 NUMERICAL RESULTS

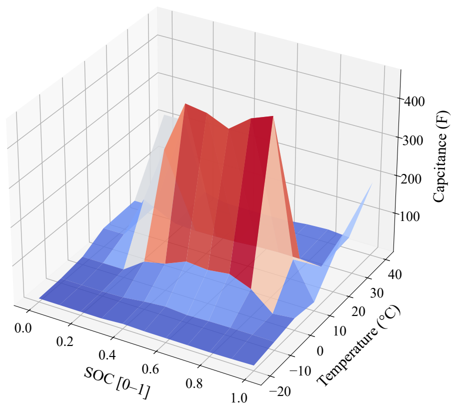
We now evaluate the performance of the DREKF algorithm for LIB SOC estimation. Numerical simulations are carried out in MATLAB and Simulink. We evaluate the DREKF on an open-source LIB dataset [33]. The battery used in this dataset is the LG 18650HG2, which is a nickel-manganese-cobalt Li-ion cell. It is specifically designed to balance high capacity (3000 mAh) with a high continuous discharge rate of up to 20A, making it suitable for EV applications. The dataset contains tests at $-20$°C, $-10$°C, 0°C, 10°C, 25°C and 40°C. HPPC tests are used to characterize the battery behaviour, and the Urban Dynamometer Driving Schedule (UDDS) tests are used to evaluate the performance of the model against baselines, namely the EKF and UKF. The UDDS dataset gives us a sample time of $\Delta t = 0.1$ s.

## 4.1 MATLAB/Simulink Implementation

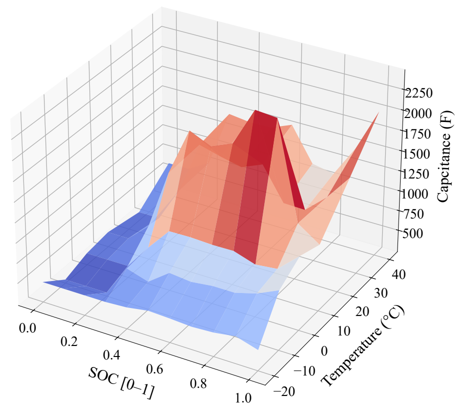
To build our simulation model, we use the MATLAB `Simscape` battery library in Simulink [92]. The library enables us to load the HPPC test directly, and an optimization function fits the parameter vector $\hat{\boldsymbol{\psi}}$ defined in (1.10). We use Simscape's `tfest` fitting function to obtain our parameters, as it provides good accuracy for both the load and relaxation proportion. When repeated over a range of temperatures, we can obtain our ECM parameters as a function of both SOC and temperature. To account for hysteresis, separate parameters are fit for charging and discharging. We showcase our resulting parameters for when the battery is discharging in Figure 4.1. The resulting parameter estimates are validated against the measured terminal voltage in the HPPC test. To this effect, we present the fitting accuracy at 40°C in Figure 4.2.

To evaluate the performance of the DREKF, the parameterized model is then represented by an ECM block provided by the `Simscape` Battery library. Within the block, the parameters are stored as a set of two-dimensional lookup tables, indexed jointly by SOC and temperature. During simulation, the ECM block returns the appropriate parameter values based on the estimated SOC and temperature at timestep $k$. We impose the simulation temperature through the block's thermal port. This architecture eliminates the need for user-defined switching logic and ensures that both the DREKF and baseline filters use state-space matrices that are uniformly updated throughout the simulation. Figure 4.3 presents a schematic which gives a high-level overview of the simulation implementation.

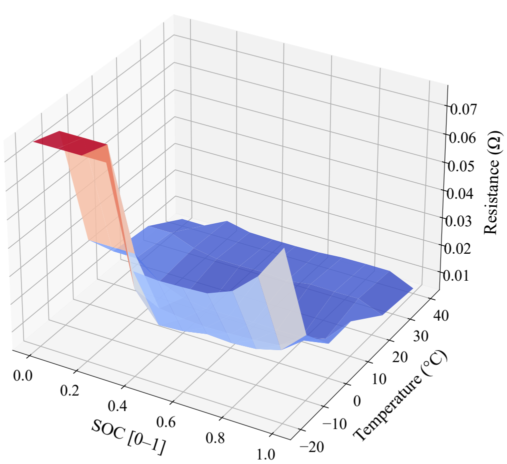
In our simulation, we assume that the battery SOH is constant at 100%. The inclusion of



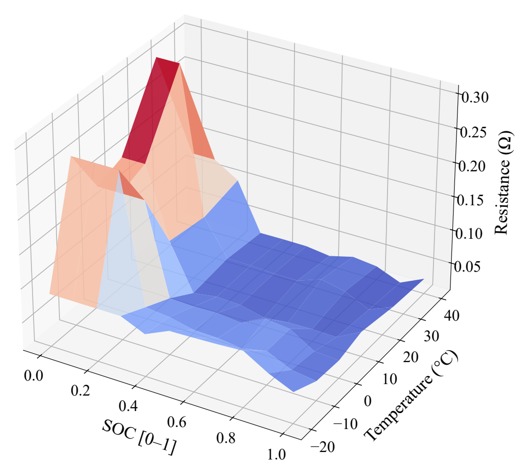

(a) $c_1(T, s)$



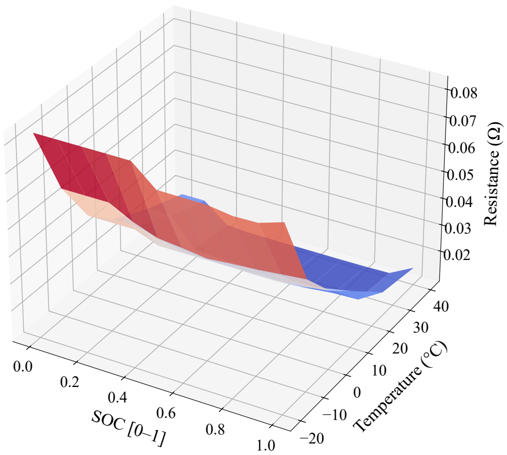

(b) $c_2(T, s)$



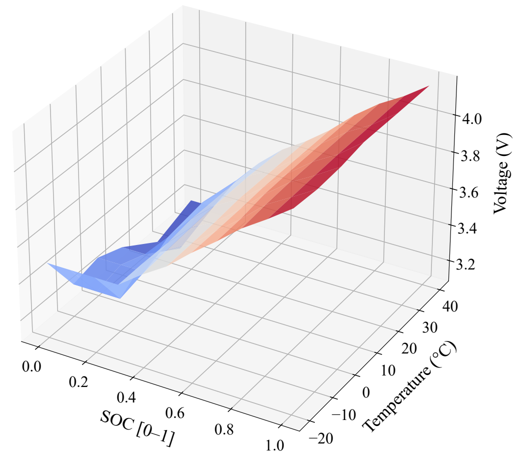

(c) $r_1(T, s)$




(d) $r_2(T, s)$




(e) $r_0(T, s)$




(f) $u_{\text{oc}}(T, s)$

Figure 4.1 Estimated parameter vector during discharge, $\hat{\boldsymbol{\psi}}$

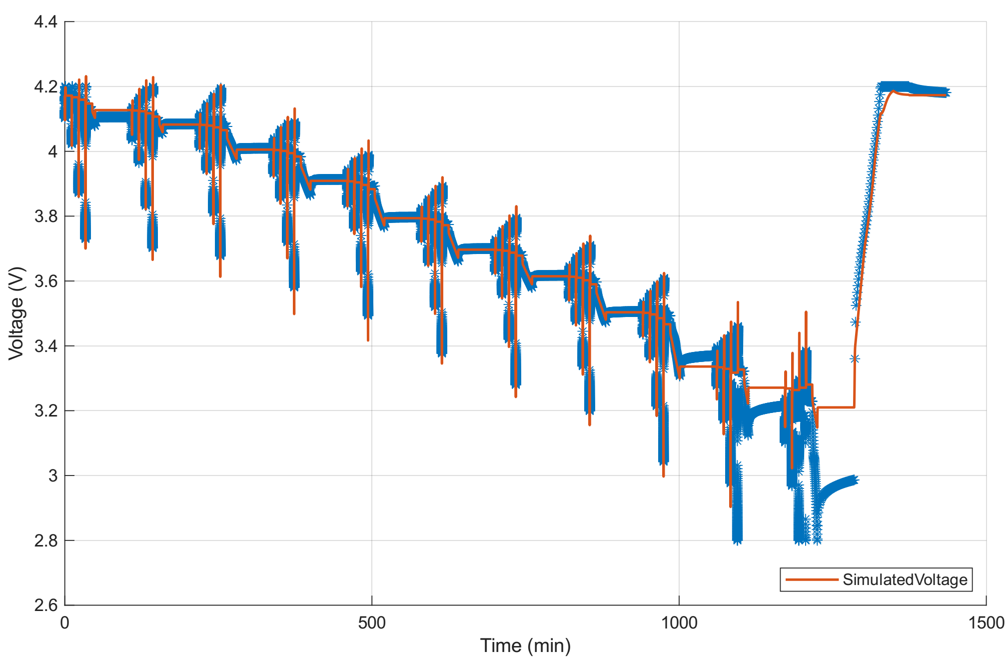



Figure 4.2 Parameter fitting using the HPPC test at 40°C

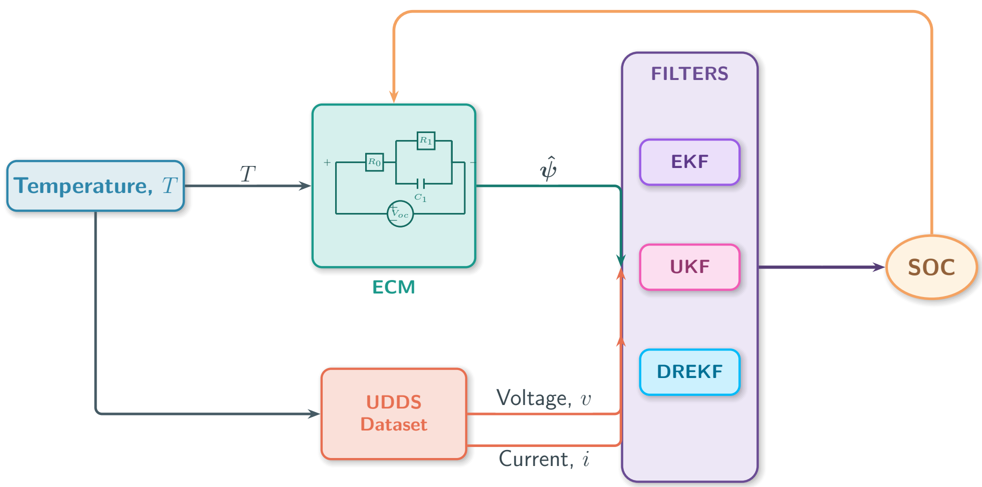



Figure 4.3 Illustration of the simulation architecture

a varying SOH is a topic for future work. The reference SOC is obtained via CC using high-precision current measurements and a known initial SOC, $x_{1,0} = 1.0$. To evaluate the correction ability of the various Kalman filters, we also assess the performance when $x_{1,0}$ does not match the true value. In this thesis, we use the initial condition $x_{1,0} = 0.75$.

For each filter, the process and measurement covariances are tuned iteratively until a stable value is obtained. For the DREKF, $\theta_{\mathbf{x}}$, $\theta_{\mathbf{v}}$, and $\epsilon$ are determined through the use of `Optuna` [93], where 100 trials are carried out to find the best performing values with respect to the root mean square error (RMSE). Through this, $\boldsymbol{\theta} = (0.01, 0.001)^\top$ is selected for our DREKF. Our full implementation is readily available on Github.

## 4.2 SOC Estimation Accuracy Under Nominal Conditions

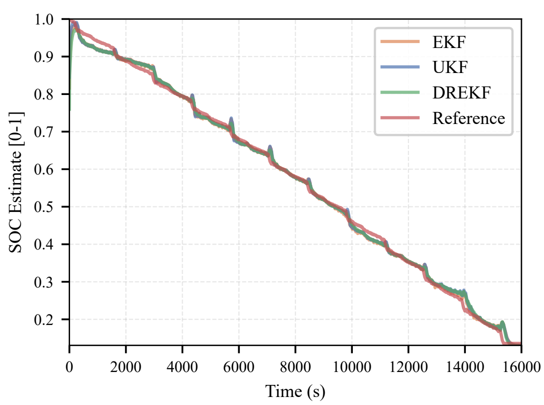
Table 4.1 reports averages for the SOC estimation error for the nominal UDDS data across temperature and initial SOC conditions. We consider the mean absolute error (MAE) and the RMSE as primary metrics for assessing the accuracy, favouring the RMSE as it gives higher weights to larger deviations from the reference value. We also look at the average 75$^{\text{th}}$ percentile and the number of outlier errors within our dataset, to evaluate the spread of the estimation errors. Outliers are defined per test as samples exceeding 1.5× the inter-quartile range of the absolute error.

When averaging the metrics across all temperatures, the proposed DREKF consistently outperforms EKF across all metrics. Under the correct initial SOC ($x_{1,0} = 1.0$), it also marginally outperforms UKF. With $x_{1,0} = 0.75$, UKF performs slightly better, partly due to its faster

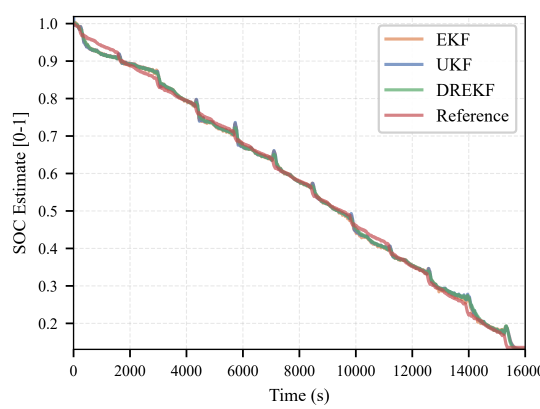
Table 4.1 Average Performance metrics across −20°C to 40°C for nominal UDDS dataset at $\epsilon = 0.01$

| $\mathbf{x}_{1,0}$ | **Method** | **MAE (%)** | **RMSE (%)** | **75$^{\mathbf{th}}$ pct.(%)** | **No. Outliers** |
|---|---|---|---|---|---|
| 1.00 | EKF | 3.14 | 4.46 | 3.85 | 3,347.67 |
| | UKF | 2.85 | 3.23 | 3.55 | 3,933.00 |
| | DREKF | **2.77** | **3.10** | **3.46** | **2,937.33** |
| 0.75 | EKF | 3.53 | 4.42 | 4.33 | 4,270.33 |
| | UKF | **2.89** | **3.30** | **3.58** | 4134.67 |
| | DREKF | 3.03 | 3.69 | **3.58** | **4015.33** |

settling time, which is illustrated in Figure 4.5. In Figure 4.4, we present a sample comparison of all the filters compared against the reference value, showing that all filters exhibit similar performance under nominal noise assumptions. Crucially, the DREKF incurs no significant performance penalty from its robustness under nominal conditions, while also preserving a settling time comparable to the baseline filters when SOC correction is required. The results also serve to validate our nominal noise assumptions for the UKF and EKF.




(a) $x_{1,0} = 0.75$




(b) $x_{1,0} = 1.0$

Figure 4.4 Comparison of filter performance at 25°C in nominal data compared to the reference at $\epsilon = 0.01$

## 4.3 Robustness against corrupted data

To evaluate the robustness of our approach, we employ a Sliced-Wasserstein (SW) outlier generation method [30] on the UDDS dataset. This method allows us to generate out-of-

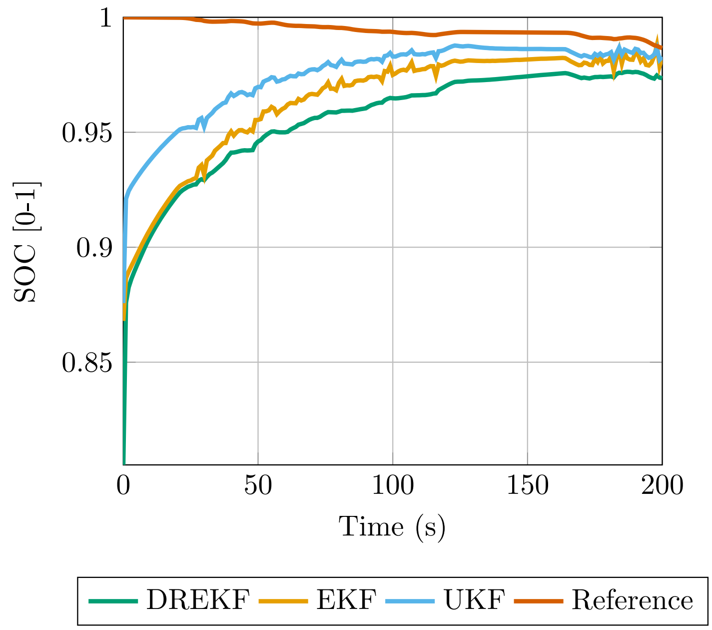
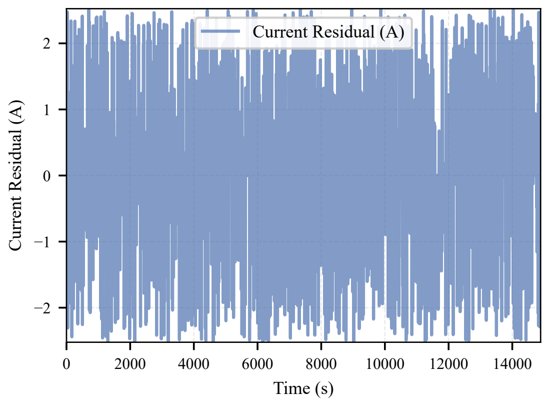



Figure 4.5 Convergence speed for the tested estimators at 25°C.

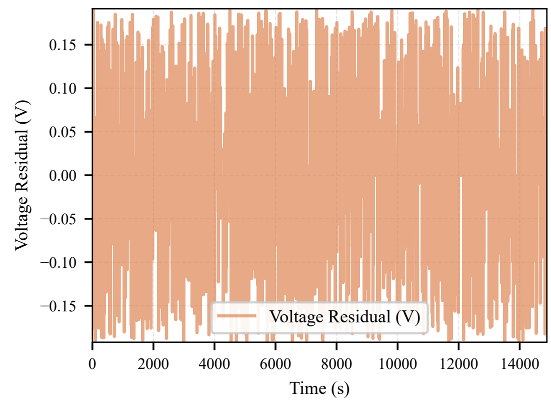
distribution outliers that can mimic non-Gaussian noise by thresholding the SW-distance between the sample and the outlier distributions. In our corrupted dataset, we artificially corrupt 1% of the sample current measurements to exceed a SW-distance threshold of 0.01, and corrupt 1% of the voltage points to exceed a SW-distance threshold of $1 \times 10^{-3}$. To provide a visual aid of this disturbance, a sample segment showing the difference between the corrupted measurement data and the nominal data is presented in Figure 4.6. To provide some intuition as to the extent of the disturbance, in certain regions of the OCV-SOC curve, a change of 0.1 V can correspond to approximately a 10% change in the SOC. The performance metrics from the test are presented in Table 4.2.

We observe in Table 4.2 that the DREKF generally outperforms the other two filters, with lower RMSE, MAE and $75^{\text{th}}$ percentile values for all temperatures, except at 10°C, where the EKF is the best estimator. However, in this case, the DREKF still outperforms the UKF, and has no outliers in comparison to the 32 for the EKF.

Across all temperatures, the primary benefit of the proposed DREKF is a reduction in the upper tail of the estimation error distribution, as evidenced by consistently lower $75^{\text{th}}$ percentile errors and fewer outliers. This reflects the fact that distributional robustness limits the sensitivity of the filter's gain to outliers, due to the consideration of the worst-case distribution. We further illustrate this in Figure 4.8, where the DREKF exhibits narrower whiskers




(a) Current Residual with SW-distance threshold of $1 \times 10^{-2}$




(b) Voltage Residual with SW-distance threshold of $1 \times 10^{-3}$

Figure 4.6 Difference between the nominal and corrupted data at 10°C

Table 4.2 Performance comparison for the 1% corrupted UDDS dataset at $x_{1,0} = 1.0$ and $\epsilon = 0.01$

| Temp | Method | MAE (%) | RMSE (%) | 75th pct. (%) | No. Outliers |
|---|---|---|---|---|---|
| −20°C | EKF | 8.64 | 9.93 | 10.40 | 3397 |
| | UKF | 7.42 | 7.89 | 8.73 | 4781 |
| | DREKF | **7.14** | **7.51** | **8.36** | **3322** |
| −10°C | EKF | 5.47 | 6.87 | 5.18 | 14435 |
| | UKF | **4.04** | 4.46 | **4.27** | 11041 |
| | DREKF | **4.04** | **4.38** | 4.40 | **8279** |
| 0°C | EKF | 4.80 | 8.24 | 3.49 | 23676 |
| | UKF | 1.89 | 2.23 | 2.70 | 256 |
| | DREKF | **1.76** | **2.09** | **2.56** | **83** |
| 10°C | EKF | 2.04 | 3.31 | 2.48 | 4268 |
| | UKF | 1.91 | 2.22 | 2.68 | 1734 |
| | DREKF | **1.87** | **2.17** | **2.63** | **1602** |
| 25°C | EKF | 1.12 | 1.46 | 1.59 | 4131 |
| | UKF | 0.95 | 1.28 | 1.43 | 3622 |
| | DREKF | **0.88** | **1.16** | **1.29** | **2673** |
| 40°C | EKF | 1.14 | 1.50 | 1.75 | 1534 |
| | UKF | 0.90 | 1.28 | 1.49 | 1969 |
| | DREKF | **0.86** | **1.20** | **1.43** | **1462** |

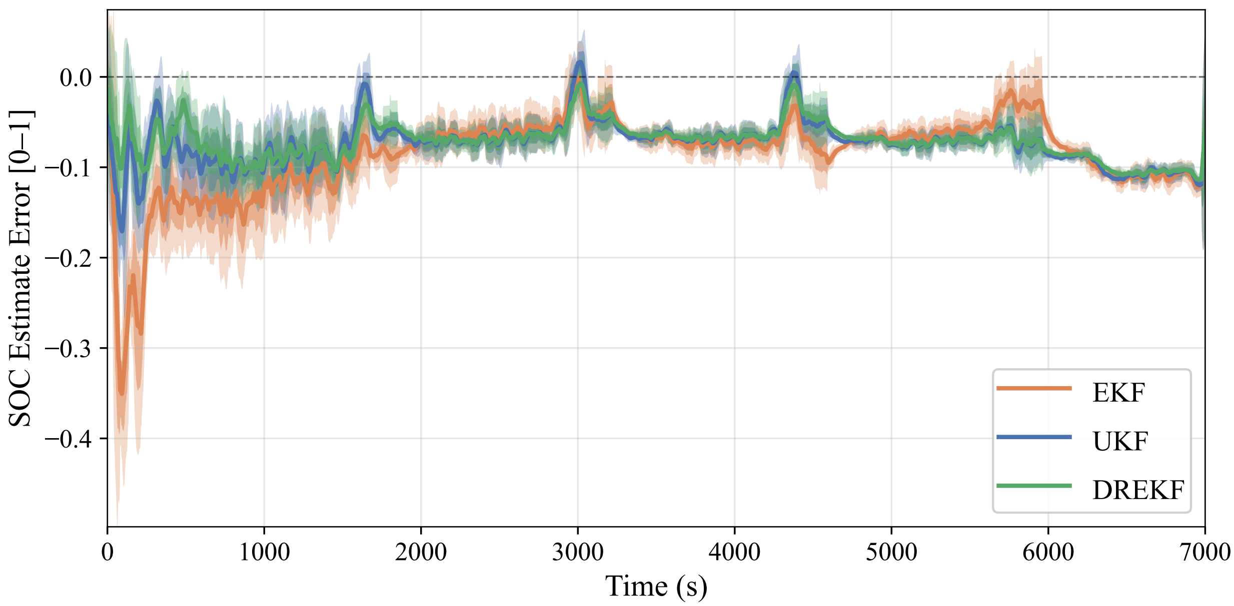
and fewer outliers relative to EKF and UKF.

We note a clear performance degradation at $-20°C$ across all estimators. Exceptionally, some outliers at this temperature correspond to anomalously low estimation errors, as shown in Figure 4.8a. At this extreme cold, the battery's internal and polarization resistances increase sharply and nonlinearly in the battery, which the parameter lookup tables fail to capture as accurately compared to the other temperatures. As such, while distributional robustness effectively mitigates the effect of non-Gaussian noise, it cannot compensate for larger, systematic model inaccuracies. Nevertheless, we observe that the DREKF is effective in reducing the maximum estimation error by approximately 5% relative to the UKF and 25% relative to the EKF.

In Figure 4.7, we display the rolling average SOC estimation error, using a sliding window of one hundred samples. We showcase the DREKF's improved disturbance rejection through the lower maximum estimation errors during periods of higher noise. At $-20°C$ in Figure 4.7a, we observe a clear bias in the error, where the estimate is consistently below the reference value, confirming that the parameter estimation does not accurately capture the battery dynamics at this temperature. In Figure 4.7b, we see that the EKF diverges in the 0°C case, indicating that the Jacobian is likely insufficient to represent the dynamics at the lower SOC. We remark that our DREKF mitigates this divergence, despite relying on the Jacobian $\mathbf{H}_k$. This suggests that the worst-case covariance obtained via the SDP provides sufficient robustness to compensate for the information loss incurred by the EKF's first-order local linearization. In Figure. 4.7a, we see that the EKF takes more aggressive correction steps within the 1$^{\text{st}}$ 200s, leading to more outlier estimations. Across Figure 4.7a- Figure 4.7d, we can see that the DREKF and UKF perform very similarly in terms of the average error, with the DREKF being slightly smoother. Overall, the DREKF provides a smoother, more consistent performance, particularly in reducing large estimation errors.

### 4.4 Computational Complexity

To satisfy real-time operation requirements, each state estimate must be computed before the next measurement is acquired. In this experiment, this corresponds to a maximum allowable computation time of 0.1 s per sample. We evaluate the impact of the tolerance parameter $\epsilon$ by carrying out simulations with varying values of $\epsilon$. We choose values within the range of $0.001 - 0.02$ V, as this roughly translates to a SOC difference of $0.1 - 2\%$ based on the battery's corresponding OCV-SOC profile. We compute the average computation times across all temperature conditions, using the corrupted UDDS dataset. All trials are conducted on a 64-bit Lenovo ThinkPad with 16GB RAM, and a 13th Gen Intel Core i7



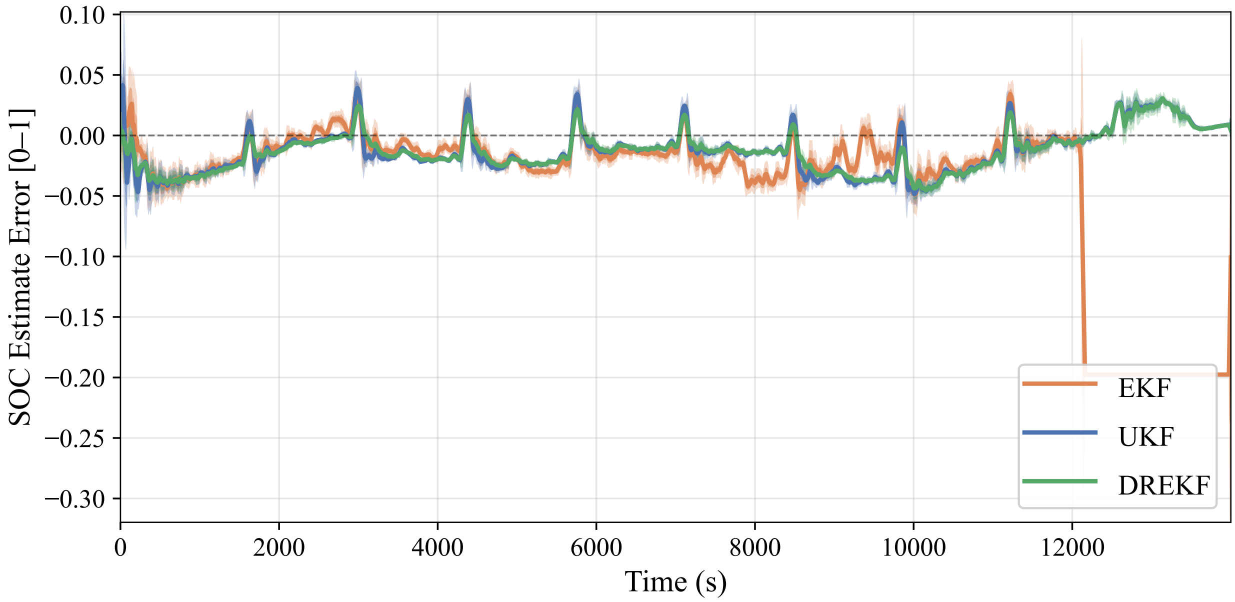

(a) Rolling average error at −20°C

processor. The `Simulink Profiler` is used to time each filter during a simulation. For the DREKF, we present the runtimes alongside the average accuracy metrics in Table 4.3.

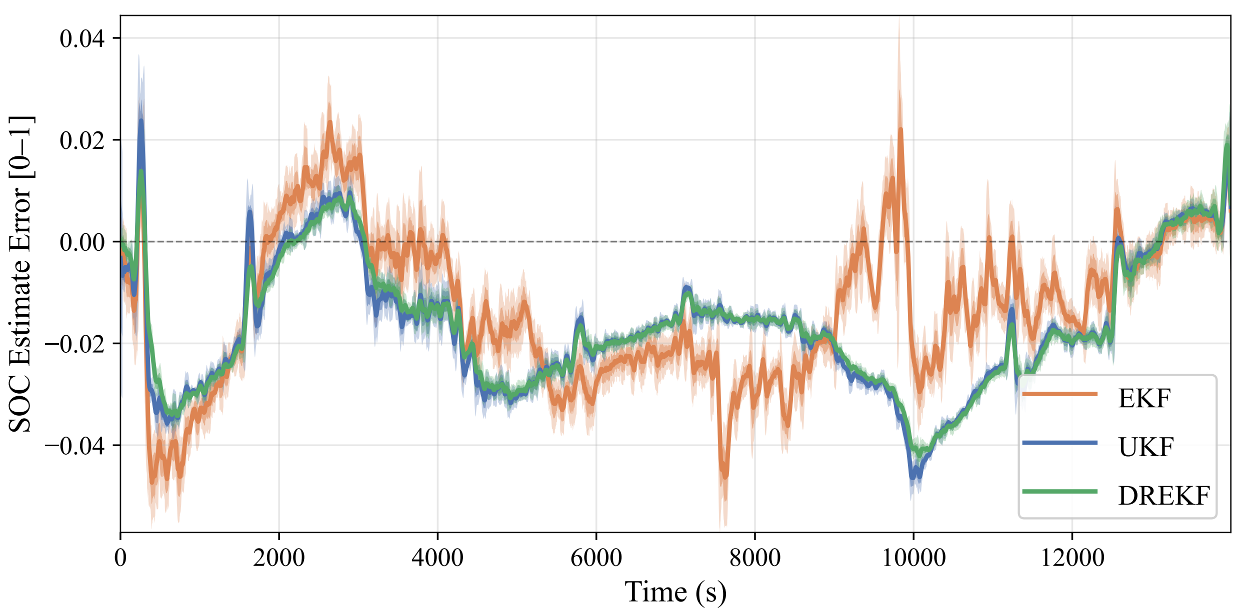
Table 4.3 Computational efficiency evaluation of the DREKF for different $\epsilon$

| $\epsilon$ | Average Computation Time (ms/sample) | MAE (%) | RMSE (%) | 75th pct. error (%) |
|---|---|---|---|---|
| 0.001 | 23.6 | 2.93 | 3.38 | 3.70 |
| 0.005 | 5.15 | 2.82 | 3.16 | 3.56 |
| 0.01 | 2.51 | 2.76 | 3.44 | 3.44 |
| 0.02 | 1.15 | 2.75 | 3.43 | 3.43 |

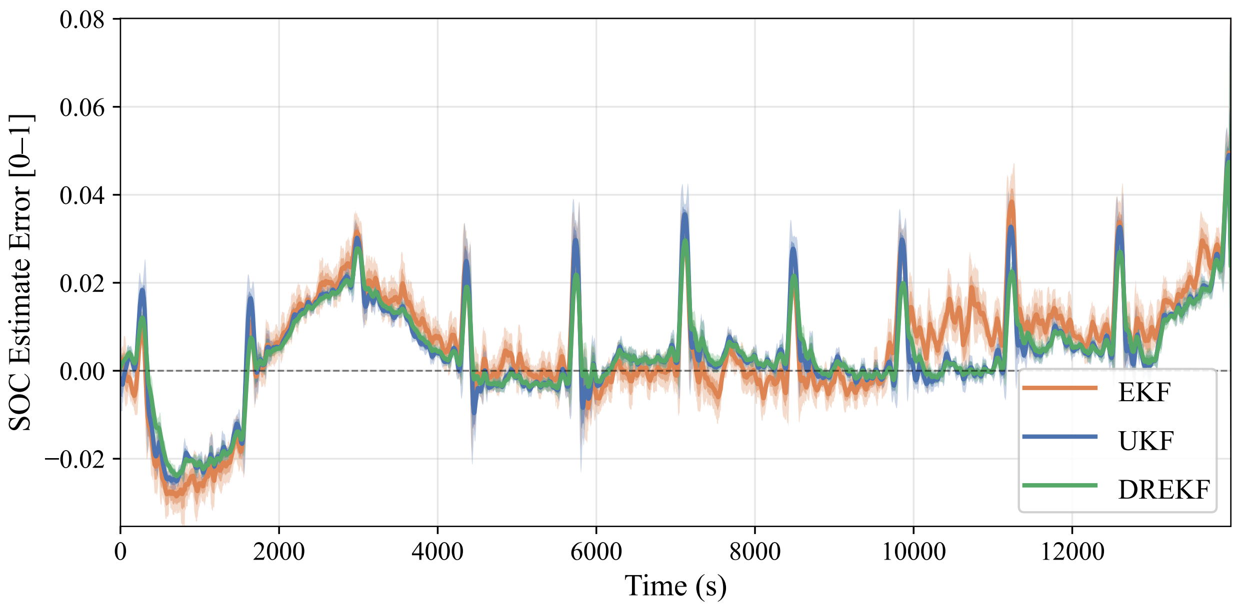
Compared to the other filters, the EKF exhibits the lowest mean execution time at 0.634 ms per sample. The UKF incurs a moderate increase in overhead, averaging 2.39 ms per sample. For the DREKF, the cost is significantly higher at the tightest tolerance of $\epsilon = 0.001$, where the SDP is recomputed at effectively every timestep. However, Table 4.3 shows that the computation cost remains comparable to the UKF at tolerances of $\epsilon = 0.005$ and 0.01, while being more efficient at $\epsilon = 0.02$. These results demonstrate that the implementation of $\epsilon$ is effective in reducing the computational demands without sacrificing the robustness of the DREKF, and the algorithm can be implemented with a marginal overhead relative to the benchmark filters. Critically, all variants operated well within the 0.1 s sampling interval, validating the real-time feasibility of the DREKF.




(b) Rolling average error at 0°C



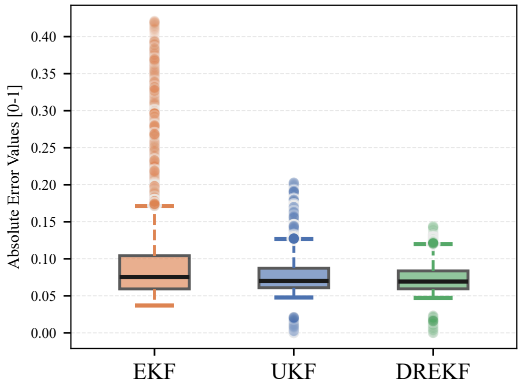

(c) Rolling average error at 10°C



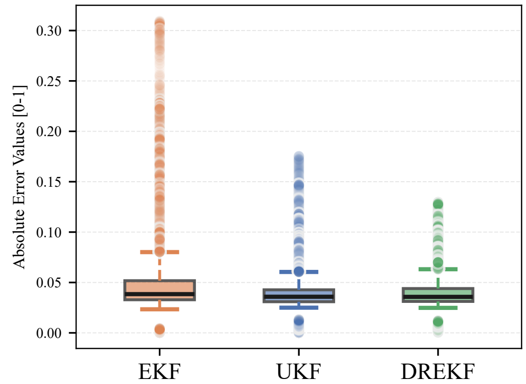

(d) Rolling average error at 40°C

Figure 4.7 Rolling average estimation error using a sliding window of 100 samples for various temperatures at $x_{1,0} = 1.0$ and $\epsilon = 0.01$

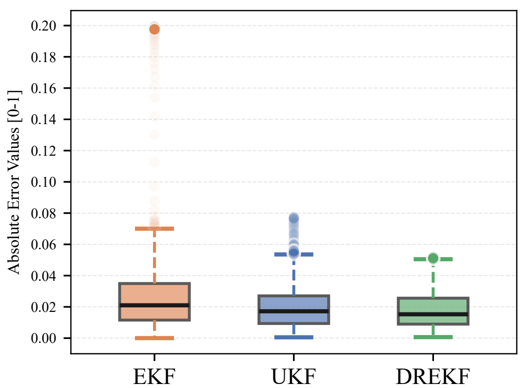
Interestingly, there is a correlation between higher tolerances and improved accuracy, which suggests that a steady-state worst-case covariance provides superior accuracy against the SW noise present in the test set. A possible explanation for this is that due to the infrequent occurrence of the corrupted measurements (1%), constant adaptation to new worst-case distributions leads to suboptimal gain fluctuations. Instead, the higher tolerance acts as a regularizer, signalling when the worst-case covariances effectively models the underlying noise statistics without over-reacting to outliers.



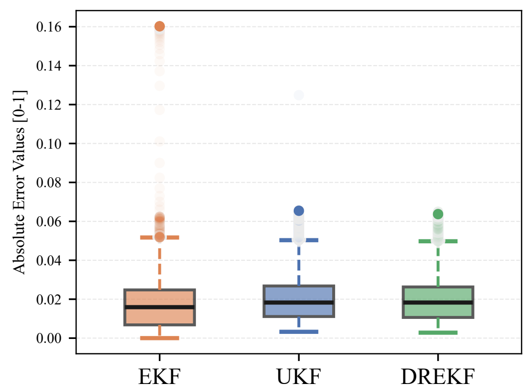

(a) SOC error distribution at $-20°C$



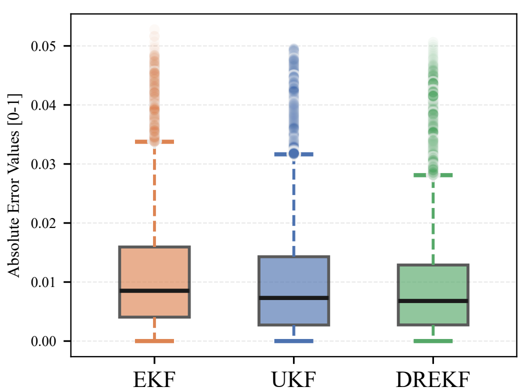

(b) SOC error distribution at $-10°C$



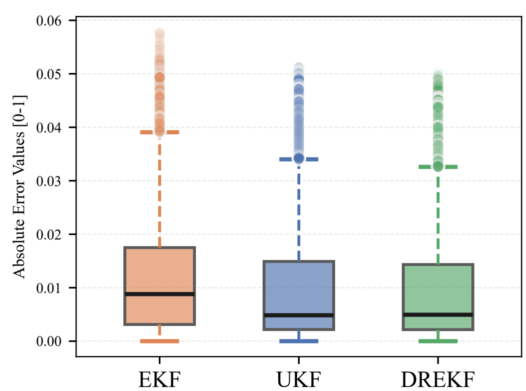

(c) SOC error distribution at $0°C$




(d) SOC error distribution at $10°C$




(e) SOC error distribution at $25°C$




(f) SOC error distribution at $40°C$

Figure 4.8 Box plot SOC error distribution of estimators at various temperatures for $x_{1,0} = 1.0$ and $\epsilon = 0.01$

# CHAPTER 5 CONCLUSION

Distributionally robust Kalman filtering represents a meaningful departure from conventional Kalman filtering approaches, which typically assume that the noise and uncertainty characterizing sensor measurements and model parameters follow a fixed, known Gaussian distribution. An electric vehicle (EV) battery management system (BMS) operates under highly variable and often unpredictable conditions, whereby the sensor noise does not necessarily conform to the nominal assumptions, and heavy-tailed noise is a possibility due to the presence of power electronics components within an EV. A distributionally robust estimator performs well not for a single assumed distribution but across an entire family of plausible distributions, meaning that its performance guarantees are preserved under adversarial deviations from the nominal distribution. In the context of a deployed EV BMS, this translates to an estimator that maintains accuracy and reliability across a broader range of real-world operating scenarios than a conventionally tuned extended Kalman filter (EKF) or unscented Kalman filter (UKF).

## 5.1 Summary of Works

This Master's thesis proposed a new distributionally robust extended Kalman filter (DREKF) for Lithium-ion battery (LIB) state-of-charge estimation. The proposed DREKF merges formulations from the distributionally robust Kalman filter (DRKF) and blends it with the EKF to render the algorithm suitable for nonlinear state estimation problems.

By incorporating Wasserstein-based ambiguity sets into the EKF estimation framework, we present a new algorithm that is able to hedge against non-Gaussian outliers and deviations from the nominal distribution estimate, while preserving computational efficiency. The DREKF provides a robustness certificate over the EKF by solving a semidefinite program (SDP), which yields the worst-case prior and posterior state covariances as well as the worst-case measurement covariance. To improve the efficiency of this algorithm for real-time estimation, we introduce an update tolerance to limit the recomputation of the SDP only to periods where the estimation value falls outside of the specified accuracy. The result of this is a computationally efficient filter that handles the nonlinear state transition and measurement functions of LIBs while providing robustness to non-Gaussian measurement outliers without being overly conservative.

The DREKF is implemented on a commonly employed driving simulation dataset, and is validated by comparing metrics against baseline filters, the EKF and the UKF, in both nominal and corrupted settings. Simulation results across multiple temperatures and operating conditions demonstrate that the DREKF achieves accuracy comparable to conventional EKF and UKF approaches under nominal conditions. Under corrupted conditions, the DREKF provides improvement compared to the conventional EKF when assessing the mean absolute error, the root mean square error, the span of the error, and the number of outliers. Compared to the UKF, the DREKF offers a more marginal gain in terms of resilience to measurement outliers. Overall, if we consider the sum of the $75^{\text{th}}$ percentile error across all of the evaluated temperatures in the corrupted case, the total of the DREKF is 20.41% lower compared to the EKF and 3.04% lower compared to the UKF. These results demonstrate that the DREKF offers a practical and meaningful improvement over classical Kalman filtering approaches for SOC estimation in LIBs, particularly in scenarios where real-world noise characteristics deviate from the nominal distributional assumptions underlying the filter design.

### 5.2 Limitations

Although the DREKF is successful in offering improved SOC estimation under uncertainty, we recognize some limitations that warrant further development of the algorithm.

1. The DREKF is still limited against low-fidelity models: Where the parameter estimation does not fully capture the transient voltage behaviour, a bias is present in the estimation error, indicating that while the DREKF is good at rejecting outliers, it cannot overcome limitations in parameter fitting.

2. Tuning of the Wasserstein radii: The DREKF requires tuning of the Wasserstein radii through a hyperparameter optimization process, in order to identify distances which increase robustness but do not detract significantly from overall accuracy.

3. To the best of our knowledge, no tight global worst-case MSE bounds exist for the DREKF. The bounded error guarantees that exist for the DRKF [80] do not directly extend to the nonlinear setting due to the linearized approximation of the measurement function. Although Figure 4.7 shows that the error bounds for the DREKF are much tighter than the error bounds of the EKF, this does not provide a generalized guarantee.

### 5.3 Future Research

The limitations mentioned above provide promising directions for future research. Given that the UKF consistently outperforms the EKF for LIB SOC estimation, a natural extension would be the development of a distributionally robust UKF framework. A theoretical motivation for this direction is that, under a moment-based characterization of distributional similarity, it has been shown that Gaussian approximation filters are distributionally robust in the sense that they employ maximum entropy prior state distributions and maximum entropy likelihood distributions [81]. However, explicitly extending this to a Wasserstein ambiguity set is nontrivial, as it requires adapting the sigma point propagation scheme to account for distributional shifts in both the mean and covariance.

Second, considering a non-constant SOH and analyzing the performance of the DREKF with battery aging would also be a productive extension of this filter. In this vein, the performance of dual-estimation DREKFs could be studied where the parameter values and battery capacity are updated simultaneously, depending on the SOC and SOH of the battery. Additionally, it would also enable a study of the degree of aging that the prescribed Wasserstein radii are able to account for. An additional improvement could be a statistically derived Wasserstein radius, based on the sampling data stored by a BMS over time, eliminating the need for tuning the Wasserstein radii. Both of these directions would provide useful contributions to advancing the applicability of the DREKF for real-time SOC estimation in a BMS .

Finally, additional work could also consider the challenges presented by pack-level SOC estimation, where cell-to-cell variability and thermal dynamics within the cells in the pack must also be considered. In this context, the DREKF may represent a particularly well-suited framework, as it could naturally accommodate a broader uncertainty set that would arise due to cell-to-cell variations and non-uniform thermal conditions.

# REFERENCES


[1] "Model Voltage Hysteresis in Battery - MATLAB & Simulink," https://www.mathworks.com/help/simscape-battery/ug/model-battery-hysteresis-voltage-example.html.

[2] A. L. C. Ferrer and A. M. T. Thomé, "Carbon Emissions in Transportation: A Synthesis Framework," *Sustainability*, vol. 15, no. 11, p. 8475, May 2023.

[3] M. M. Islam, X. Zhong, Z. Sun, H. Xiong, and W. Hu, "Real-time Frequency Regulation Using Aggregated Electric Vehicles in Smart Grid," *Computers & Industrial Engineering*, vol. 134, pp. 11–26, Aug. 2019.

[4] "Trends in batteries – Global EV Outlook 2023 – Analysis," https://www.iea.org/reports/global-ev-outlook-2023/trends-in-batteries.

[5] J. Tao, S. Wang, W. Cao, P. Takyi-Aninakwa, C. Fernandez, and J. M. Guerrero, "A Comprehensive Review of State of Charge and State of Health Estimation for Lithium-ion Battery Energy Storage Systems," *Ionics*, vol. 30, no. 10, pp. 5903–5927, Oct. 2024.

[6] F. Alanazi, "Electric Vehicles: Benefits, Challenges, and Potential Solutions for Widespread Adaptation," *Applied Sciences*, vol. 13, no. 10, May 2023.

[7] O. Tremblay, L.-A. Dessaint, and A.-I. Dekkiche, "A generic battery model for the dynamic simulation of hybrid electric vehicles," in *2007 IEEE Vehicle Power and Propulsion Conference*, vol. 284. Ieee, 2007.

[8] O. Tremblay and L.-A. Dessaint, "Experimental validation of a battery dynamic model for EV applications," *World electric vehicle journal*, vol. 3, no. 2, pp. 289–298, 2009.

[9] V. Lucaferri, M. Quercio, A. Laudani, and F. Riganti Fulginei, "A Review on Battery Model-Based and Data-Driven Methods for Battery Management Systems," *Energies*, vol. 16, no. 23, p. 7807, Jan. 2023.

[10] K. Movassagh, A. Raihan, B. Balasingam, and K. Pattipati, "A Critical Look at Coulomb Counting Approach for State of Charge Estimation in Batteries," *Energies*, vol. 14, no. 14, p. 4074, Jan. 2021.

[11] Q. Chen, F. Zhang, L. Su, B. Lin, S. Chen, and Y. Zhang, “State Parameter Fusion Estimation for Intelligent Vehicles Based on IMM-MCCKF,” *Applied Sciences*, vol. 14, no. 11, May 2024.

[12] D. Qi, J. Feng, X. Ni, and L. Wang, “Maximum Correntropy Extended Kalman Filter for Vehicle State Observation,” *International Journal of Automotive Technology*, vol. 24, no. 2, pp. 377–388, Apr. 2023.

[13] W. Song, R. Liu, X. Jin, and W. Guo, “SOC Estimation for Lithium-ion Batteries Based on Weighted Multi-Innovation Sage–Husa Adaptive EKF,” *Energies*, vol. 18, no. 16, Aug. 2025.

[14] M. Wu, L. Qin, and G. Wu, “State of Charge Estimation of Power Lithium-ion Battery Based on an Adaptive Time Scale Dual Extended Kalman filter,” *Journal of Energy Storage*, vol. 39, p. 102535, Jul. 2021.

[15] D. Kuhn, S. Shafiee, and W. Wiesemann, “Distributionally Robust Optimization,” *Acta Numerica*, vol. 34, pp. 579–804, Jul. 2025.

[16] B. Balasingam, *Robust Battery Management System Design with MATLAB*. Norwood, MA, USA: Artech House Publishers, Jun. 2023.

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

[18] M. A. A. Mohamed, T. F. Yu, G. Ramsden, J. Marco, and T. Grandjean, “Advancements in Parameter Estimation Techniques for 1RC and 2RC Equivalent Circuit Models of Lithium-ion Batteries: A Comprehensive Review,” *Journal of Energy Storage*, vol. 113, p. 115581, Mar. 2025.

[19] J. Peng, J. Meng, J. Wu, Z. Deng, M. Lin, S. Mao, and D.-I. Stroe, “A comprehensive overview and comparison of parameter benchmark methods for lithium-ion battery application,” *Journal of Energy Storage*, vol. 71, p. 108197, Nov. 2023.

[20] Q. Song, Y. Mi, and W. Lai, “A Novel Variable Forgetting Factor Recursive Least Square Algorithm to Improve the Anti-Interference Ability of Battery Model Parameters Identification,” *IEEE Access*, vol. 7, pp. 61 548–61 557, 2019.

[21] X. Sun, J. Ji, B. Ren, C. Xie, and D. Yan, “Adaptive Forgetting Factor Recursive Least Square Algorithm for Online Identification of Equivalent Circuit Model Parameters of a Lithium-ion Battery,” *Energies*, vol. 12, no. 12, p. 2242, Jan. 2019.

[22] X. Peng, J. Yin, L. Sun, Z. Ye, and T. Wei, “Parameter Identification of Lithium-ion Battery Equivalent Circuit Model Based on Limited Memory Recursive Least Squares Algorithm with Variable Forgetting Factor,” *Journal of Physics: Conference Series*, vol. 2246, no. 1, p. 012090, Apr. 2022.

[23] C. Ma, C. Wu, L. Wang, X. Chen, L. Liu, Y. Wu, and J. Ye, “A Review of Parameter Identification and State of Power Estimation Methods for Lithium-ion Batteries,” *Processes*, vol. 12, no. 10, p. 2166, Oct. 2024.

[24] Z. Wang, D. T. Gladwin, M. J. Smith, and S. Haass, “Practical State Estimation Using Kalman Filter Methods for Large-Scale Battery Systems,” *Applied Energy*, vol. 294, p. 117022, Jul. 2021.

[25] X. Lin, Y. Tang, J. Ren, and Y. Wei, “State of Charge Estimation with the Adaptive Unscented Kalman Filter Based on an Accurate Equivalent Circuit Model,” *Journal of Energy Storage*, vol. 41, p. 102840, Sep. 2021.

[26] E. Wan and R. Van Der Merwe, “The unscented Kalman filter for nonlinear estimation,” in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium*, Oct. 2000, pp. 153–158.

[27] S. J. Julier and J. K. Uhlmann, “New extension of the Kalman filter to nonlinear systems,” in *Signal Processing, Sensor Fusion, and Target Recognition VI*, vol. 3068. SPIE, Jul. 1997, pp. 182–193.

[28] S. Küçükyavuz and R. Jiang, “Chance-Constrained Optimization under Limited Distributional Information: A Review of Reformulations Based on Sampling and Distributional Robustness,” *EURO Journal on Computational Optimization*, vol. 10, p. 100030, Jan. 2022.

[29] P. Mohajerin Esfahani and D. Kuhn, “Data-Driven Distributionally Robust Optimization Using the Wasserstein Metric: Performance Guarantees and Tractable Reformulations,” *Mathematical Programming*, vol. 171, no. 1, pp. 115–166, Sep. 2018.

[30] J. Pallage and A. Lesage-Landry, “Wasserstein Distributionally Robust Shallow Convex Neural Networks,” *INFORMS Journal on Optimization*, vol. 8, no. 1, pp. 61–93, Jan. 2026.

[31] S. Shafieezadeh Abadeh, V. A. Nguyen, D. Kuhn, and P. M. Mohajerin Esfahani, “Wasserstein Distributionally Robust Kalman Filtering,” in *Advances in Neural Infor-*

*mation Processing Systems*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., vol. 31. Curran Associates, Inc., 2018.

[32] M. Jang, A. Hakobyan, and I. Yang, "On the Steady-State Distributionally Robust Kalman Filter," in *2025 IEEE 64th Conference on Decision and Control (CDC)*, Dec. 2025, pp. 2385–2392.

[33] P. Kollmeyer, C. Vidal, M. Naguib, and M. Skells, "LG 18650HG2 Li-ion Battery Data and Example Deep Neural Network xEV SOC Estimator Script," *Mendeley Data*, vol. 3, Mar. 2020.

[34] S.-L. Wang, X. Xiong, C.-Y. Zou, L. Chen, C. Jiang, Y.-X. Xie, and D.-I. Stroe, "An improved coulomb counting method based on dual open-circuit voltage and real-time evaluation of battery dischargeable capacity considering temperature and battery aging," *International Journal of Energy Research*, vol. 45, no. 12, pp. 17 609–17 621, 2021.

[35] A. Jokar, B. Rajabloo, M. Désilets, and M. Lacroix, "Review of Simplified Pseudo-two-Dimensional Models of Lithium-ion batteries," *Journal of Power Sources*, vol. 327, pp. 44–55, Sep. 2016.

[36] D. Zhang, S. Dey, L. D. Couto, and S. J. Moura, "Battery Adaptive Observer for a Single-Particle Model With Intercalation-Induced Stress," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 4, pp. 1363–1377, Jul. 2020.

[37] S. Moura, F. Bribiesca Argomedo, R. Klein, A. Mirtabatabaei, and M. Krstic, "Battery State Estimation for a Single Particle Model With Electrolyte Dynamics," *IEEE Transactions on Control Systems Technology*, vol. 25, pp. 1–16, Jun. 2016.

[38] M. Doyle, T. F. Fuller, and J. Newman, "Modeling of Galvanostatic Charge and Discharge of the Lithium/Polymer/Insertion Cell," *Journal of The Electrochemical Society*, vol. 140, no. 6, p. 1526, Jun. 1993.

[39] S. J. Moura, N. A. Chaturvedi, and M. Krstic, "PDE estimation techniques for advanced battery management systems — Part I: SOC estimation," in *2012 American Control Conference (ACC)*, Jun. 2012, pp. 559–565.

[40] D. Zhang, B. N. Popov, and R. E. White, "Modeling Lithium Intercalation of a Single Spinel Particle under Potentiodynamic Control," *Journal of The Electrochemical Society*, vol. 147, no. 3, p. 831, Mar. 2000.

[41] M. Guo, G. Sikha, and R. E. White, “Single-Particle Model for a Lithium-Ion Cell: Thermal Behavior,” *Journal of The Electrochemical Society*, vol. 158, no. 2, p. A122, Dec. 2010.

[42] W. Zschiebsch, J. Merker, and R. Bohm, “The Mathematical Pitfalls of the Original Single Particle Model,” in *Proceedings of the 2023 6th International Conference on Mathematics and Statistics*, ser. ICoMS ’23. New York, NY, USA: Association for Computing Machinery, Dec. 2023, pp. 19–25.

[43] Z. Huang, L. D. Couto, C. Dangwal, S. Xiao, W. Lv, D. Zhang, and S. J. Moura, “On Electrochemical Model-Based State Estimation for Lithium–Sulfur Batteries,” *IEEE Transactions on Control Systems Technology*, vol. 32, no. 3, pp. 849–861, May 2024.

[44] A. Zilali, M. Adda, K. Ziane, and M. Berger, “Machine learning-based state of charge estimation: A comparison between CatBoost model and C-BLSTM-AE model,” *Machine Learning with Applications*, vol. 20, p. 100629, Jun. 2025.

[45] T. Talluri, H. T. Chung, and K. Shin, “Study of Battery State-of-charge Estimation with kNN Machine Learning Method,” *IEIE Transactions on Smart Processing & Computing*, vol. 10, no. 6, pp. 496–504, Dec. 2021.

[46] J. Dou, H. Ma, Y. Zhang, S. Wang, Y. Ye, S. Li, and L. Hu, “Extreme Learning Machine Model for State of Charge Estimation of Lithium-ion Battery Using Salp Swarm Algorithm,” *Journal of Energy Storage*, vol. 52, p. 104996, Aug. 2022.

[47] Y. Zhou, S. Wang, Y. Xie, T. Zhu, and C. Fernandez, “An improved particle swarm optimization-least squares support vector machine-unscented Kalman filtering algorithm on SOC estimation of lithium-ion battery,” *International Journal of Green Energy*, vol. 21, no. 2, pp. 376–386, Jan. 2024.

[48] S.-B. Kim and S.-H. Lee, “Application of Deep Learning Techniques for the State of Charge Prediction of Lithium-Ion Batteries,” *Applied Sciences*, vol. 14, no. 17, p. 8077, Jan. 2024.

[49] M. A. Hannan, M. S. H. Lipu, A. Hussain, M. H. Saad, and A. Ayob, “Neural Network Approach for Estimating State of Charge of Lithium-Ion Battery Using Backtracking Search Algorithm,” *IEEE Access*, vol. 6, pp. 10 069–10 079, 2018.

[50] M. H. Lipu, M. A. Hannan, A. Hussain, M. Saad, A. Ayob, and K. M. Muttaqi, “Lithium-ion Battery State of Charge Estimation Method Using Optimized Deep Recurrent Neural

Network Algorithm," in *2019 IEEE Industry Applications Society Annual Meeting*, Sep. 2019, pp. 1–9.

[51] K. L. Wong, M. Bosello, R. Tse, C. Falcomer, C. Rossi, and G. Pau, "Li-Ion Batteries State-of-Charge Estimation Using Deep LSTM at Various Battery Specifications and Discharge Cycles," in *Proceedings of the Conference on Information Technology for Social Good*, ser. GoodIT '21. New York, NY, USA: Association for Computing Machinery, Sep. 2021, pp. 85–90.

[52] Q. Wen, View Profile, T. Zhou, View Profile, C. Zhang, View Profile, W. Chen, View Profile, Z. Ma, View Profile, J. Yan, View Profile, L. Sun, and View Profile, "Transformers in time series," in *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, ser. Guide Proceedings, Aug. 2023, pp. 6778–6786.

[53] M. Yılmaz, E. Çinar, and A. Yazıcı, "A Transformer-Based Model for State of Charge Estimation of Electric Vehicle Batteries," *IEEE Access*, vol. 13, pp. 33 035–33 048, 2025.

[54] L. Li, Y. Li, R. Mao, L. Li, W. Hua, and J. Zhang, "Remaining Useful Life Prediction for Lithium-Ion Batteries With a Hybrid Model Based on TCN-GRU-DNN and Dual Attention Mechanism," *IEEE Transactions on Transportation Electrification*, vol. 9, no. 3, pp. 4726–4740, Sep. 2023.

[55] J. Zhao, X. Han, Y. Wu, Z. Wang, and A. F. Burke, "Opportunities and challenges in transformer neural networks for battery state estimation: Charge, health, lifetime, and safety," *Journal of Energy Chemistry*, vol. 102, pp. 463–496, Mar. 2025.

[56] K.-c. Bae, S.-c. Choi, J.-h. Kim, C.-y. Won, and Y.-c. Jung, "LiFePO4 dynamic battery modeling for battery simulator," in *2014 IEEE International Conference on Industrial Technology (ICIT)*, Feb. 2014, pp. 354–358.

[57] N. Campagna, V. Castiglia, R. Miceli, R. A. Mastromauro, C. Spataro, M. Trapanese, and F. Viola, "Battery Models for Battery Powered Applications: A Comparative Study," *Energies*, vol. 13, no. 16, Aug. 2020.

[58] T. Kim, A. Adhikaree, D. Kang, M. Kim, C.-Y. Oh, and J. Baek, "A real-time condition monitoring for lithium-ion batteries using a low-price microcontroller," in *2017 IEEE Energy Conversion Congress and Exposition (ECCE)*, Oct. 2017, pp. 5248–5253.

[59] E. D. Maria, M. Dallapiccola, D. Aloisio, G. Brunaccini, F. Sergi, D. Moser, and G. Barchi, "A data-driven equivalent circuit model's parameter estimation method ap-

plied to Lithium-Titanate battery," *Journal of Energy Storage*, vol. 74, p. 109497, Dec. 2023.

[60] Y. Chu, J. Li, L. Li, and Y. Qiang, "Parameter identification and SOC estimation of lithium-ion battery based on AGCOA optimization," in *2022 34th Chinese Control and Decision Conference (CCDC)*, Aug. 2022, pp. 5964–5968.

[61] H. Dai, T. Xu, L. Zhu, X. Wei, and Z. Sun, "Adaptive model parameter identification for large capacity Li-ion batteries on separated time scales," *Applied Energy*, vol. 184, pp. 119–131, Dec. 2016.

[62] "Simscape," The Mathworks, Inc., Natick, Massachusetts, United States, 2025.

[63] G. L. Plett, "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, Aug. 2004.

[64] J. Yun, Y. Choi, J. Lee, S. Choi, and C. Shin, "State-of-Charge Estimation Method for Lithium-ion Batteries Using Extended Kalman Filter With Adaptive Battery Parameters," *IEEE Access*, vol. 11, pp. 90 901–90 915, 2023.

[65] Z. Liu, Z. Zhao, Y. Qiu, B. Jing, C. Yang, and H. Wu, "Enhanced State of Charge Estimation for Li-ion Batteries Through Adaptive Maximum Correntropy Kalman Filter with Open Circuit Voltage Correction," *Energy*, vol. 283, p. 128738, Nov. 2023.

[66] H. Lv, Y. Liao, C. Zhao, X. Shang, and F. Zhang, "State of charge estimation of lithium-titanate battery based on multi-model extended Kalman filter considering temperature and current rate," *Journal of Energy Storage*, vol. 77, p. 109890, Jan. 2024.

[67] A. Almagbile, J. Wang, and W. Ding, "Evaluating the Performances of Adaptive Kalman Filter Methods in GPS/INS Integration," *Journal of Global Positioning Systems*, vol. 9, Jun. 2010.

[68] S. Peng, A. Zhang, D. Liu, M. Cheng, J. Kan, and M. Pecht, "State-of-Charge Estimation of Lithium-ion Batteries Based on Dual-Coefficient Tracking Improved Square-Root Unscented Kalman Filter," *Batteries*, vol. 9, no. 8, Jul. 2023.

[69] Z. Zhou, M. Zhan, B. Wu, G. Xu, X. Zhang, J. Cheng, and M. Gao, "A novel adaptive unscented kalman filter algorithm for SOC estimation to reduce the sensitivity of attenuation coefficient," *Energy*, vol. 307, p. 132598, Oct. 2024.

[70] R. Van der Merwe and E. Wan, “The square-root unscented Kalman filter for state and parameter-estimation,” in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings*, vol. 6, May 2001, pp. 3461–3464 vol.6.

[71] L. Chen, X. Wu, A. M. Lopes, L. Yin, and P. Li, “Adaptive State of Charge Estimation of Lithium-ion Batteries Based on Square-Root Unscented Kalman Filter,” *Energy*, vol. 252, p. 123972, Aug. 2022.

[72] O. Rezaei, A. Rahdan, S. Sardari, M. Dahmardeh, and Z. Wang, “A fuzzy robust two-stage unscented Kalman filter method for uncertainty and state of charge estimation of lithium-ion batteries,” *Journal of Energy Storage*, vol. 68, p. 107883, Sep. 2023.

[73] J. Doyle, “Robust and optimal control,” in *Proceedings of 35th IEEE Conference on Decision and Control*, vol. 2, Dec. 1996, pp. 1595–1598 vol.2.

[74] M. Zarif, M. Charkhgard, and A. Alfi, “Hybrid state of charge estimation for lithium-ion batteries: Design and implementation,” *IET Power Electronics*, vol. 7, pp. 2758–2764, Nov. 2014.

[75] B. Xia, Z. Zhang, Z. Lao, W. Wang, W. Sun, Y. Lai, and M. Wang, “Strong Tracking of a H-Infinity Filter in Lithium-Ion Battery State of Charge Estimation,” *Energies*, vol. 11, p. 1481, Jun. 2018.

[76] N. Y. Ko and T. G. Kim, “Comparison of Kalman Filter and Particle Filter Used for Localization of an Underwater Vehicle,” in *2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, Nov. 2012, pp. 350–352.

[77] B. Liu, Y. Xu, H. Zhang, J. Zhang, F. Yang, Y. Wang, H. Yang, Z. Liu, and H. Zheng, “State of Charge Estimation of Supercapacitor Under Different Temperatures Using Particle Filter Algorithm Based on Fractional-Order Model,” *Journal of The Electrochemical Society*, vol. 170, no. 9, p. 090541, Sep. 2023.

[78] Y. Xu, X. Chen, H. Zhang, F. Yang, L. Tong, Y. Yang, D. Yan, A. Yang, M. Yu, Z. Liu, and Y. Wang, “Online Identification of Battery Model Parameters and Joint State of Charge and State of Health Estimation using Dual Particle Filter Algorithms,” *International Journal of Energy Research*, vol. 46, no. 14, pp. 19 615–19 652, 2022.

[79] C. Kuptametee and N. Aunsri, “A review of resampling techniques in particle filtering framework,” *Measurement*, vol. 193, p. 110836, Apr. 2022.

[80] J. Hajar, T. Kargin, V. Malik, and B. Hassibi, "Distributionally Robust Kalman Filtering over an Infinite-Horizon," in *2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2025, pp. 1–5.

[81] S. Wang, "Distributionally Robust State Estimation for Nonlinear Systems," *IEEE Transactions on Signal Processing*, vol. 70, pp. 4408–4423, 2022.

[82] Y. Wu, D. Bai, K. Zhang, Y. Li, and F. Yang, "Advancements in the Estimation of the State of Charge of Lithium-ion Battery: A Comprehensive Review of Traditional and Deep Learning Approaches," *Journal of Materials Informatics*, vol. 5, no. 2, Feb. 2025.

[83] "AI-Enhanced Battery Management Systems for EVs," https://resources.altium.com/p/ai-enhanced-battery-management-systems-evs, Jul. 2024.

[84] E. Chemali and M. Preindl, "Neural-network state-of-charge estimation," US Patent US20 190 157 891A1, May, 2019.

[85] M. Ul Hassan, S. Saha, M. E. Haque, S. Islam, A. Mahmud, and N. Mendis, "A Comprehensive Review of Battery State of Charge Estimation Techniques," *Sustainable Energy Technologies and Assessments*, vol. 54, p. 102801, Dec. 2022.

[86] M. Gelbrich, "On a Formula for the L2 Wasserstein Metric between Measures on Euclidean and Hilbert Spaces," *Mathematische Nachrichten*, vol. 147, no. 1, pp. 185–203, 1990.

[87] V. A. Nguyen, S. Shafieezadeh-Abadeh, D. Kuhn, and P. Mohajerin Esfahani, "Bridging Bayesian and Minimax Mean Square Error Estimation via Wasserstein Distributionally Robust Optimization," *Mathematics of Operations Research*, vol. 48, no. 1, pp. 1–37, Feb. 2023.

[88] A. Hakobyan and I. Yang, "Wasserstein Distributionally Robust Control of Partially Observable Linear Stochastic Systems," *IEEE Transactions on Automatic Control*, vol. 69, no. 9, pp. 6121–6136, Sep. 2024.

[89] D. Kuhn, P. M. Esfahani, V. A. Nguyen, and S. Shafieezadeh-Abadeh, "Wasserstein Distributionally Robust Optimization: Theory and Applications in Machine Learning," in *Operations Research & Management Science in the Age of Analytics*, ser. INFORMS TutORials in Operations Research. INFORMS, Oct. 2019, ch. 6, pp. 130–166.

[90] J. Huang, D. Shi, and T. Chen, "Event-triggered robust state estimation for systems with unknown exogenous inputs," *Automatica*, vol. 122, p. 109248, Dec. 2020.

[91] ——, "Robust event-triggered state estimation: A risk-sensitive approach," *Automatica*, vol. 99, pp. 253–265, Jan. 2019.

[92] "Simscape Battery Documentation," https://www.mathworks.com/help/simscape-battery/index.html.

[93] T. Akiba, View Profile, S. Sano, View Profile, T. Yanase, View Profile, T. Ohta, View Profile, M. Koyama, and View Profile, "Optuna: A Next-Generation Hyperparameter Optimization Framework," in *The 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, ser. ACM Conferences, Jul. 2019, pp. 2623–2631.