PHYSICS-INSPIRED AND CONTROL-ORIENTED MODELING OF LITHIUM BATTERIES FOR ACCURATE STATE-OF-CHARGE PREDICTION AND FAST-CHARGING

A Thesis
Submitted to
the Temple University Graduate Board

In Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

---

by
Renato Rodriguez Nunez
Diploma Date December 2024

Examining Committee Members:

Damoon Soudbakhsh, Ph.D., Advisor, Department of Mechanical Engineering

Philip Dames, Ph.D., Department of Mechanical Engineering

Fei Ren, Ph.D., Department of Mechanical Engineering

Joseph Picone, Ph.D., Department of Electrical and Computer Engineering

Yan Wang, Ph.D., Nikola Corporation

ABSTRACT

This dissertation presents a physics-inspired, data-driven framework for lithium-ion battery modeling and control, designed to enhance state prediction accuracy and optimize fast-charging strategies across diverse operating conditions. It addresses key challenges in battery management systems for critical applications such as electric vehicles, focusing on improving performance, adaptability, and safety under extreme temperatures, high charge/discharge currents, and varying states of charge. Conventional methods often fail to fully utilize battery capacity, particularly in low-temperature conditions, resulting in reduced performance and limited range. The research proposes a novel methodology combining data-driven techniques with physics-based insights to overcome these issues. A key innovation is the development of *PhITEDD* (Physics-Informed Temperature Dependent Explicit Data-Driven), a framework that enhances model accuracy, interpretability, and generalizability while reducing reliance on proprietary knowledge.

The *PhITEDD* framework combines physics-inspired features that connect the model to underlying battery processes, a Monte Carlo search algorithm for exploring extensive feature spaces, and an automated hyperparameter tuning mechanism. This approach strikes an optimal balance between model accuracy and complexity by quantifying individual feature contributions, enabling robust state prediction across a wide range of operating conditions. Its digital twin for state-of-charge forecasting achieves prediction errors below 1% using experimental drive cycle data, while maintaining high accuracy for unseen aggressive drive cycles and varying operating conditions.

A key innovation of the framework is its temperature-dependent recalibration method, which adjusts model coefficients to optimize performance under new operating conditions, ensuring consistency across the full SOC (0%-100%) and temperature (-20°C to 40°C) range. The framework also investigates the impact of data sampling rates on model accuracy, providing practical guidelines for optimization. These advancements collectively enhance the interpretability, efficiency, and practicality of lithium-ion battery models, supporting improved battery utilization and extended lifespan.

The dissertation also addresses the fast-charging optimization problem using a direct data-driven control method. This strategy learns the battery's Jacobian from input/output data to optimize the charging current profile, minimizing charging time while adhering to safety constraints such as maximum cell temperature and voltage. The data was generated using a full-order electrochemical Doyle-Fuller-Newman model integrated with a thermal model. The optimal solution comprises a hybrid charging strategy that charges a 5Ah NMC-811 cylindrical cell 66% faster than the standard CCCV method, while ensuring safety limits like 4.2V and 57°C. This approach closely aligns with actual battery mechanisms.

By improving efficiency and safety, the proposed methodology has significant implications for the performance and adaptability of batteries in electric vehicles and other critical applications. The work reduces dependency on proprietary models, enhancing the accessibility and applicability of battery modeling tools. In summary, this dissertation advances Li-ion battery research by integrating physics-informed and data-driven methods, resulting in innovative modeling and control strategies for accurate SOC prediction and optimal fast charging under complex operating conditions.

KEYWORDS: Data-driven modeling, Data-driven control, Hyperparameter autotunning, Monte Carlo Library Search, Jacobian Learning, Li-ion Batteries, SOC modeling, and Electric Vehicles.

DEDICATION

This thesis is dedicated to my parents, Jesus Rodriguez and Dora Nunez, whose unwavering support and love have always been my foundation. To my beloved wife, Danielle Rodriguez, for her endless patience, encouragement, and understanding throughout this journey. And to my precious children, Addison and Elijah Rodriguez, who inspire me to strive for excellence.

Thank you for believing in me.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Damoon Soudbakhsh. His insightful mentorship, thoughtful guidance, and unwavering support have been invaluable to me. I am particularly grateful for his remarkable ability to distill complex ideas, identify technical challenges, and propose innovative solutions. Above all, Dr. Soudbakhsh believed in me, and for that, I am profoundly thankful.

I would also like to extend my sincere thanks to my committee members, Dr. Phillip Dames and Dr. Fei Ren, for their thoughtful questions and invaluable feedback. Their advice has been instrumental in laying a strong foundation for the work presented in this dissertation.

I am deeply appreciative of all my colleagues in the Dynamical Systems Laboratory (DSLab) for their mentorship and friendship throughout my doctoral studies. Special thanks go to Omidreza Ahmadzadeh and Mohsen Derakhshan for their support, technical discussions, expertise in battery experimental facilities, and most importantly, their camaraderie. This work would not have been possible without their contributions.

Lastly, I would like to express my immense gratitude to the Office of Naval Research (ONR grant number N000142312612), Ford Motor Company (University Research Project (URP) grant), Temple University Graduate School, and the College of Engineering for their financial support, which made it possible for me to pursue my research endeavors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The market share of electric vehicles (EVs), measured by the number of vehicles on the road, has experienced steady growth over the past decade [1], as shown in Fig 1.1. This growth has been driven by advancements in battery technology, stricter emissions regulations to reduce environmental pollution, and increasing consumer demand for cleaner and more sustainable transportation options [1]. The shift towards electrification is further bolstered by the rising awareness of climate change and the adverse environmental impacts of internal combustion engine (ICE) vehicles, such as greenhouse gas emissions and air quality deterioration. As EV technology matures, battery costs have dramatically decreased, making EVs more affordable for the average consumer while offering higher energy densities, longer driving ranges, and faster charging times. This synergy between technological progress and market forces has paved the way for the rapid expansion of EVs globally.

This upward trend in EV adoption is expected to continue as governments worldwide implement ambitious policy frameworks to accelerate the transition to a low-carbon economy. Countries like Norway, Germany, China, and the United States have introduced a variety of incentives, such as tax credits, rebates, and infrastructure investments, to promote EV adoption [1]. Many of these nations have also set aggressive timelines to phase out the production and sale of ICE vehicles, pushing automakers to pivot towards electric powertrains. Global initiatives have also encouraged countries to decarbonize their transportation sectors, cementing EVs as a cornerstone of future mobility.

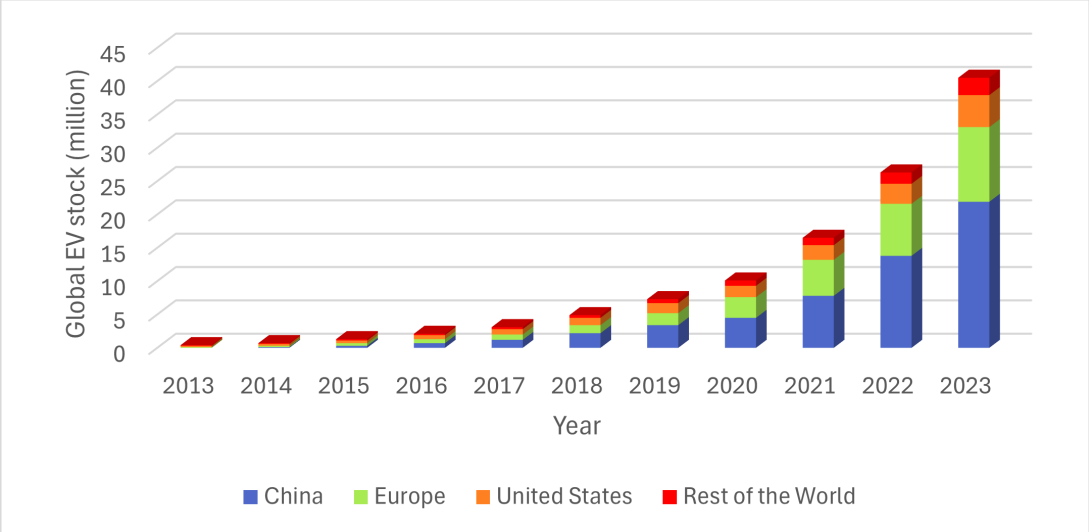

Figure 1.1. Global Electric Vehicle Stock Growth (2013–2023).

As EV adoption continues to grow globally, so does the demand for optimized energy storage systems that provide reliable and consistent performance. Central to this challenge is the need to manage and enhance the efficiency of lithium-ion batteries, the dominant energy storage solution in modern EVs. Lithium-ion batteries offer several advantages, including high energy density, low self-discharge rates, and a decreasing cost curve. However, they also present complexities, as their performance is affected by factors such as temperature fluctuations, charge/discharge cycles, and aging, all of which can reduce battery capacity over time. While alternative energy storage solutions, such as solid-state batteries, offer higher safety, energy density, and durability, their current high manufacturing costs and scalability challenges limit their widespread use in EVs.

1.1 Battery Management

Efficient energy management, is a critical factor in ensuring that EVs can meet the demands of both consumers and regulatory bodies [1–3]. The ability

to optimize the use of stored energy within the battery pack is directly linked to several key performance indicators. These include the vehicle's driving range, acceleration, and overall energy efficiency. Beyond performance, enhancing operational safety is another priority. As EVs rely on high-voltage battery systems, managing thermal stability and avoiding overcharging or deep discharging are essential to prevent catastrophic failures, such as thermal runaway or battery fires. Moreover, extending the lifespan of the battery, which is typically one of the most expensive components in an EV, is crucial for reducing total ownership costs and improving the economic viability of electric transportation.

Therefore, the continued development and refinement of battery management systems (BMS) is paramount. BMS technology is central to monitoring battery health, ensuring safety, and maximizing efficiency by optimizing charging and discharging processes. As EV adoption increases and market demands grow, advancements in BMS technology will play a vital role in making electric mobility a mainstream reality.

To achieve these objectives, modern battery management systems have evolved into highly sophisticated systems, far surpassing the basic control mechanisms of early designs [2, 4]. Initially, BMS was primarily focused on simple tasks like monitoring battery voltage and temperature, protecting against overcharging and deep discharging, and balancing cells within the battery pack. However, as electric vehicles (EVs) have become more complex and expectations around performance, safety, and longevity have increased, the role of BMS has expanded significantly. Today's BMS must provide not only basic protections but also an in-depth understanding of the battery's internal states, enabling precise management of its dynamic behaviors in real-time [5].

Modern BMS are expected to operate under a wide range of conditions and stresses, from high-power acceleration and regenerative braking to extreme temperature variations [6]. This requires a sophisticated understanding of the battery's electrochemical processes, which are inherently nonlinear and influenced by numerous factors, including aging, temperature, charge/discharge rates, and the mechanical stress on cells during operation. Advanced modeling techniques, such as physics-based models, machine learning algorithms, and data-driven approaches, have become essential tools for BMS to accurately predict and manage these internal states under dynamic operating conditions [7].

In addition to real-time monitoring, modern BMS are also responsible for implementing predictive maintenance strategies. By continuously analyzing the battery's historical performance data and applying predictive algorithms, BMS can identify early signs of potential failures or capacity degradation. This allows for proactive interventions, such as adjusting charging strategies or scheduling maintenance, to mitigate issues before they escalate, thus enhancing both the safety and longevity of the battery [6].

One of the primary challenges in modern battery management systems (BMS) is accurately estimating critical battery parameters, such as state of charge (SOC), state of health (SOH), and the thermal distribution within the cells [2,5,8]. These internal states are crucial for optimizing battery performance, maintaining safety, and prolonging the battery's operational lifespan. Unlike earlier systems that relied on basic monitoring, today's BMS are embedded with advanced algorithms, sensor networks, and computational models that enable continuous monitoring and precise assessment of these parameters. By interpreting data such as current, voltage, temperature, and external factors like driving patterns and environmental

conditions, modern BMS offer a far more comprehensive understanding of the battery's health, behavior, and potential issues [9].

The state of charge is among the most critical parameters in a battery system, indicating the remaining energy as a percentage of total capacity; much like a fuel gauge in gasoline-powered vehicles [10, 11]. Accurate SOC estimation is essential not only for efficient energy management but also for providing real-time feedback to the vehicle's control system and driver, allowing precise calculation of the available driving range. Additionally, it is fundamental for maintaining safe battery operation, helping to prevent overcharging and deep discharging, both of which can shorten battery life and, in extreme cases, pose risks like thermal runaway. This insight is vital for optimizing vehicle performance, planning charging stops, and ensuring a seamless, reliable electric vehicle experience [6]. Finally, precise SOC estimation facilitates energy-efficient charging strategies, a growing priority with the rise of fast-charging technology.

Determining SOC in real-time, however, is a complex problem due to the battery's nonlinear electrochemical behavior, which varies with factors such as temperature, current load, and aging. Unlike fuel level in gasoline engines, SOC cannot be directly measured. Instead, it must be estimated from observable signals, such as electrical current, voltage, and temperature [12–14]. This estimation process is critical to the functionality of the BMS, as even small inaccuracies in SOC prediction can lead to significant performance issues or safety risks [15, 16]. Therefore, developing robust algorithms for accurate SOC prediction based on these measurable signals is crucial for the next generation of EVs.

1.2 Fast Charging

The rapid advancement and widespread adoption of fast-charging technology have introduced new complexities and demands on battery management systems. Fast charging is one of the most highly sought-after features in electric vehicles, as it drastically reduces the time required to recharge the battery, enhancing convenience for long-distance travel and improving the overall user experience.

The demand for faster charging places significant pressure on battery management systems to carefully manage and balance multiple factors, including temperature regulation, charge rate optimization, and safety protocols, while also preserving battery longevity. To meet these elevated demands, BMS must now be highly adaptive, effectively managing risks like excessive heat generation that could trigger thermal runaway and jeopardize battery safety. Thermal runaway is one of the most critical risks associated with fast charging, where excessive heat generation can trigger a chain reaction that leads to catastrophic failure, including the possibility of fire or explosion [17]. For reliable operation, BMS must also counteract various degradation effects; such as electrolyte decomposition, electrode breakdown, increased growth rate of the solid electrolyte interface (SEI) layer, and lithium plating; all of which can significantly impact battery performance and lifespan [18].

Electrolyte decomposition, for instance, can lead to gas formation and increased internal resistance, while electrode decomposition may cause irreversible changes in material structure, both impacting the cell's ability to maintain stable performance over time. Lithium plating, which occurs when lithium ions deposit as metallic lithium on the anode, not only reduces charge capacity but also raises the risk of short circuits. Additionally, the accelerated growth of the solid electrolyte interface (SEI) layer increases resistance, further contributing to capacity fade [19].

This capacity fade, a critical concern, progressively reduces the total energy the battery can store, diminishing its overall effectiveness in high-demand applications such as electric vehicles.

To address these challenges, modern BMS require advanced control algorithms that dynamically adjust charging parameters based on critical indicators such as temperature, voltage, state of charge, and state of health. These intelligent systems aim to predict aging patterns and optimize charging strategies to extend battery life while maintaining efficiency and safety. Furthermore, to mitigate the risk of thermal runaway, thermal management becomes a vital aspect of BMS, especially during high-power charging sessions where the risk of overheating is amplified [20].

Modern BMS must, therefore, be equipped with advanced thermal management systems that not only monitor temperature at various points within the battery pack but also dynamically adjust charging parameters to mitigate excessive heat buildup. This involves integrating heat dissipation strategies, such as liquid cooling systems, and real-time control algorithms that can reduce charging rates when thermal thresholds are approached, thereby ensuring the battery remains within safe operating temperatures [21].

In addition to thermal management, the electrical stress imposed by rapid charging can lead to uneven current distribution across the battery cells, exacerbating imbalances between individual cells over time. This can result in some cells being overcharged or discharged more rapidly than others, increasing the likelihood of premature failure. Modern BMS must account for these variations by employing cell balancing techniques to ensure uniform charging and discharging across the entire pack, thereby optimizing the overall system performance and enhancing safety [6, 22].

Furthermore, fast charging introduces unique challenges for accurately estimating the battery's state of charge (SOC). Traditional SOC estimation methods, which are often calibrated for slower, more stable charging cycles, may struggle to maintain accuracy at higher charging speeds. Rapid fluctuations in current and voltage during fast charging can disrupt the precision of these estimations, leading to potential issues such as undercharging, overcharging, and accelerated battery degradation [23].

To address these challenges, modern BMS must leverage more sophisticated models that can dynamically adapt to the nonlinear and temperature-dependent conditions of fast charging. These advanced algorithms must be able to not only improve SOC prediction under such conditions but also integrate real-time data from multiple sensors to account for temperature and other key factors. By providing more reliable SOC estimates, these models can enhance the BMS's ability to maintain optimal battery performance, minimize excessive degradation, and ensure safety throughout the charging process.

As fast-charging technology and electric vehicle (EV) advancements progress, the role of sophisticated battery management systems (BMS) is becoming increasingly crucial in supporting sustainable and reliable electric mobility. To enable fast-charging capabilities, BMS must be more intelligent and responsive than ever, actively managing trade-offs among charging speed, thermal stability, and battery longevity. Key features like advanced thermal management, precise cell balancing, and adaptive control systems are essential for achieving safe, high-performance fast charging in EVs without compromising long-term durability.

Modern BMS provide automakers with real-time insights into a battery's internal states, which is vital for optimizing energy storage systems and meeting the growing demands for performance, safety, and efficiency. To accomplish this,

these systems require advanced battery modeling and control methods capable of real-time adjustments based on specific operating conditions. By delivering precise and responsive charging profiles, these next-generation BMS not only enhance the convenience and appeal of EVs but also play a pivotal role in propelling the global shift toward electric mobility. As EV technology continues to evolve, BMS will be key in maximizing energy utilization, extending battery life, and supporting emerging innovations like fast charging, thereby driving the future of sustainable transportation.

1.3 Proposed Work and Contributions

This work aims to address the growing challenges in battery management systems by introducing PhITEDD (Physics-Informed Temperature-Dependent Explicit Data-Driven), a novel digital twin framework specifically designed for accurate, real-time state-of-charge forecasting. PhITEDD combines the power of physics-based insights with data-driven techniques to enhance the prediction and modeling of SOC dynamics, which is critical for optimizing battery performance and ensuring safe operation across a variety of operating environments.

PhITEDD is constructed through an explicit data-driven approach, allowing it to efficiently model complex battery behaviors and predict SOC under a broad spectrum of conditions, including extreme temperatures and varying charge/discharge cycles. The framework is highly adaptable and capable of accommodating the nonlinearities and complexities associated with aggressive charging and discharging scenarios where traditional SOC estimation methods often struggle.

By integrating temperature dependence into its formulation, PhITEDD accounts for the substantial impact of temperature fluctuations on battery perfor-

mance, improving its ability to predict SOC accurately even under thermal stress. This capability is especially crucial for electric vehicles and other high-demand applications where battery temperature can vary significantly during operation.

Furthermore, PhITEDD is designed to offer real-time estimation, ensuring that the battery remains within safe operational limits throughout its charge and discharge cycles. This proactive SOC forecasting helps mitigate risks associated with overcharging, deep discharging, or thermal runaway, ultimately extending the lifespan and improving the reliability of the battery. By leveraging physics-based terms with data-driven algorithms, PhITEDD provides a robust and highly precise solution for managing charge levels, enhancing the overall safety and performance of energy storage systems.

Additionally, this research introduces a novel fast-charging strategy aimed at optimizing the electrical current input while minimizing the degradation of lithium-ion batteries. As fast-charging technology becomes a critical factor in the widespread adoption of electric vehicles, it is essential to develop charging methods that not only reduce charging times but also protect the longevity and performance of the battery. The proposed strategy achieves this by integrating advanced learning algorithms and optimization techniques, which dynamically adjust the charging profiles based on feedback from the battery's state of charge, temperature, and voltage response.

One of the key innovations of this strategy is its ability to adapt to changing battery conditions, allowing for precise control of charging current and voltage at every stage of the charging process. This dynamic adjustment helps reduce the thermal stress typically associated with fast charging, which can lead to overheating and increase the risk of thermal runaway or other safety concerns. By optimizing current input, the strategy mitigates several aging mechanisms

that contribute to battery degradation, including lithium plating, excessive solid electrolyte interface growth, and internal resistance increase, all of which significantly impact the battery's performance and lifespan.

In addition to reducing thermal stress and preventing these degradation mechanisms, the fast-charging strategy also improves overall charging efficiency. By strategically managing the charging rates and adapting to varying conditions, it accelerates charging times while maintaining the battery within safe operational limits. This not only enables faster recharging of EVs but also ensures that the battery's health is preserved, helping reduce the frequency and severity of capacity fade over time. As a result, this research contributes to the development of more sustainable and efficient fast-charging solutions, supporting the broader goal of accelerating the transition to electric mobility while addressing the growing demand for faster and safer charging methods.

By combining advanced SOC prediction with an optimized fast-charging method, this work offers a comprehensive solution to some of the most pressing challenges hindering the widespread adoption of electric vehicle technology. Furthermore, this research paves the way for safer, more efficient, and longer-lasting energy storage systems, supporting the future of electric mobility.

The contributions of this research are threefold: a) contributions to the field of data-driven modeling, b) contributions to the field of control-oriented battery dynamics modeling, and c) contributions to developing safe and efficient battery-fast charging strategies.

(a) Enhancements to Generic Data-driven Modeling (Sparse Identification of Nonlinear Dynamics (SINDy)):

- Library Term Quantification: By quantifying the importance of individual library terms, we enable the discovery of parsimonious and

generalizable models, reducing reliance on proprietary knowledge or detailed internal parameters of individual cells.

- Monte Carlo Search for Nonlinear Terms: A Monte Carlo search efficiently explores the high-dimensional feature space, improving the representation of complex, nonlinear behaviors in LiBs.
- Optimization of Data Sampling Rates: We examine the impact of data sampling rates on model accuracy and optimize them to ensure improved performance.
- Hyperparameter Auto-Tuning: A hyperparameter auto-tuning approach identifies optimal coefficients, balancing model accuracy and complexity while minimizing the need for manual tuning.

(b) Control-oriented battery dynamics modeling:

- Accurate Reduced-Order SOC Model: We develop a reduced-order SOC model for LiBs based on individual cell current and voltage data. This model provides valuable insights for BMS, enhancing overall battery performance and safety.
- Physics-Inspired Model Initialization: Enhanced the machine learning library with physics-informed features derived from the electrochemical Doyle-Fuller-Newman (DFN) model, capturing fundamental battery processes such as intercalation, electrochemical kinetics, and diffusion. This approach replaced generic nonlinear terms, significantly improving both interpretability and accuracy.
- Generalizable Model Framework: Created a physics-informed modeling approach that improves generalizability and reduces the risk of fitting

into an incorrect nonlinear model, making it robust for diverse battery applications.

- Real-Time Suitability: Designed a model that is interpretable and suitable for real-time analysis, supporting dynamic operational environments.
- Control-Oriented Battery Modeling: Designed control-oriented models that approximate the battery dynamics into a computationally manageable form, ensuring applicability for tasks like state of charge forecasting or fast charging control.

(c) Physical Constraints and Fast Charging Optimization: Physical constraints are integrated into the optimization problem, enabling direct data-driven control for fast charging. The approach is validated through a high-fidelity, full-order electrochemical-thermal battery simulator, ensuring optimal solutions for real-world applications.

This approach is versatile and can be adapted to a wide range of energy storage systems and battery types, accommodating different cell chemistries, form factors, and operational constraints. Additionally, it is well-suited for constraint-based optimization of other complex dynamical systems, providing a valuable framework for advancing machine learning models across diverse applications.

2. LITERATURE REVIEW

Lithium-ion (Li-ion) batteries have emerged as the preferred energy storage solution across various applications, from personal electronics to electric vehicles (EVs). These batteries exhibit complex internal processes, including diffusion, intercalation, and electrochemical kinetics, which are highly sensitive to operating conditions such as temperature, state of charge (SOC), and aging. Accurately modeling these dependencies is critical for developing advanced battery management systems (BMS) to optimize performance, ensure safety, and prolong battery life. Therefore, the literature extensively explores studies focused on addressing key challenges in BMS by improving their adaptability, efficiency, and performance under varying operating conditions (e.g., temperatures, currents, and SOC levels) through the development of advanced battery models, through physics-based, data-driven, and hybrid modeling frameworks. Additionally, as the EV market expands, fueled by advancements in battery technology, stricter emissions regulations, and increasing demand for sustainable transportation, fast-charging technology has become increasingly important. While fast charging enhances convenience and user experience, it also introduces new challenges for BMS. Hence, development of optimal fast-charging strategies for critical real-time applications, aimed at optimizing performance through minimization of the battery charge time while delivering improved safety, efficiency, and battery longevity are investigated through simple (model-free) and complex (model-based) approaches. To address these challenges, the development of optimal fast-charging strategies for critical real-time applications has been a key area of investigation. These strategies

aim to minimize battery charge time while ensuring improved safety, efficiency, and longevity. Both simple (model-free) and complex (model-based) approaches are explored to optimize performance effectively. The remainder of this chapter provides a detailed review of the literature on physics-based, data-driven, and hybrid battery models, as well as passive and active fast-charging strategies, presented in separate sections.

2.1 Battery Modeling

In recent years, battery modeling and characterization have received extensive focus [12, 14, 24–47]. A primary goal of these models is to generate accurate estimations of essential battery states, such as voltage and state-of-charge, which are critical for optimal electric vehicle (EV) performance, safety, efficiency, and battery lifespan. For example, accurate SOC estimation ensures the battery is neither overcharged nor over-discharged, preventing degradation and reducing safety risks like cell damage or thermal runaway.

To address SOC estimation, researchers have explored a variety of modeling approaches. Physical models, such as the Doyle-Fuller-Newman (DFN) model [41], use first-principles to detail internal electrochemical processes, achieving high accuracy [27,33]. Direct measurement methods like open-circuit voltage (OCV) [48–50] and Coulomb counting [51–53] offer simpler solutions but may lack precision under dynamic operating conditions. Equivalent circuit models (ECMs) provide a practical approach by representing the battery as circuits of voltage sources and passive components like resistors and capacitors [36, 54, 55]. Furthermore, advanced machine learning techniques, such as artificial neural networks (ANNs) [56–59], leverage machine learning to derive battery models directly from empirical data [31, 60].

As EV adoption grows, the need for accurate, real-time estimations of these battery states becomes even more pressing. These models are integral to battery management systems, supporting decision-making that maximizes battery performance while safeguarding long-term health and safety. Achieving models that balance accuracy and computational efficiency remains crucial for effective real-time SOC estimation, essential for the demands of modern EV applications.

The following sections present a thorough review of the literature on battery modeling and SOC estimation methods. The acronyms used throughout this section are summarized in Table 2.1.

2.1.1 Physics-based Models

Physical models, including the single-particle model (SPM) [61] and the Doyle-Fuller-Newman (DFN) or pseudo-two-dimensional (P2D) model [27, 62, 63], are grounded in first-principle methods, offering a comprehensive and detailed representation of the electrochemical processes occurring within lithium-ion cells. These models are integral for understanding and predicting battery behavior under a wide range of conditions, as they explicitly model fundamental phenomena occurring within the cell structure.

For instance, the DFN model (Fig. 2.1) captures the intricate, microscopic dynamics of the battery by employing partial differential equations (PDEs) to describe key processes such as ionic diffusion, lithium-ion intercalation and deintercalation, electrolyte concentration gradients, and electrochemical reaction kinetics [27, 33]. Specifically, the DFN model treats the electrodes as porous structures, accounting for the movement of lithium ions in both the electrolyte and solid active materials within the electrodes. Ionic diffusion within the electrodes and electrolyte is modeled by Fick's law, which governs how ions spread from

regions of high to low concentration. Additionally, the model describes the electrochemical reaction kinetics at the interface between the solid electrode particles and the electrolyte using the Butler-Volmer equation, which relates the current density to the overpotential and reaction rate constants. This detailed approach allows for an accurate representation of how lithium ions are transported through the cell, stored within the electrode particles, and ultimately contribute to the battery's overall voltage and capacity.

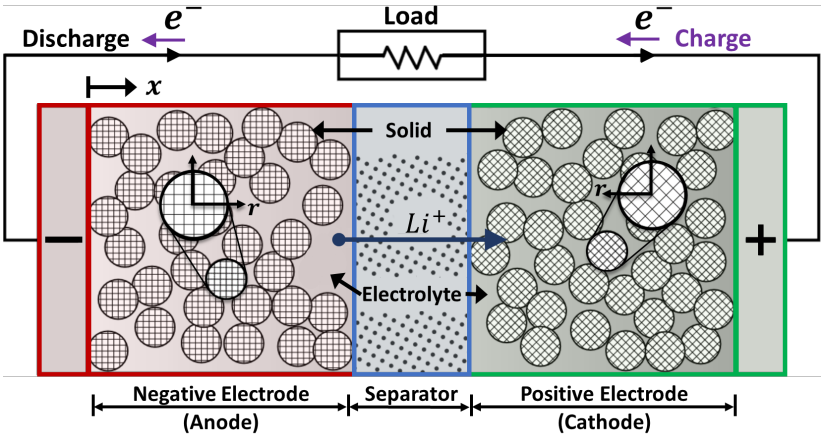

Figure 2.1. Diagram of Li-ion Battery.

Furthermore, the DFN model includes terms for the electric potential distribution across the cell, allowing for the calculation of the cell's open-circuit voltage and internal resistances under varying states of charge and temperature. This level of detail provides insights into the potential drops across the solid and liquid phases, helping understand the cell's efficiency and power capabilities under high-current applications. Additionally, by modeling electrolyte concentration gradients, the DFN model helps predict conditions that may lead to lithium plating or accelerated aging, which are crucial for high-power applications like fast charging.

Despite the high accuracy and predictive capabilities of models like DFN, they are computationally demanding. Solving the PDEs that describe solid-state diffusion and electrolyte transport requires substantial computational resources, especially when simulating batteries with large electrode surface areas or high energy densities. Each layer of the cell (anode, separator, cathode) has unique transport properties and reaction characteristics, which must be discretized and solved across multiple nodes, further increasing the computational complexity. As a result, using these models for real-time battery management in electric vehicles or other applications is challenging without significant computational simplifications.

Furthermore, accurate parameterization is crucial for the reliability of physical battery models, as it necessitates comprehensive experimental data. For instance, the DFN model alone requires over 30 parameters, specific to each cell type, which encompass material properties, physical dimensions, and electrochemical characteristics [25, 64, 65]. Developing these models involves complex and invasive experimental methods, including cell dissection, electrochemical impedance spectroscopy, and pulse-relaxation testing, to accurately identify internal battery parameters [27,41,62]. These parameters; such as the diffusion coefficient, reaction rate constants, and transference number; not only vary between different cell chemistries but also fluctuate with changes in operational conditions and battery state of health. As the battery cycles and undergoes degradation, these values shift, adding further complexity to the modeling process and challenging the maintenance of accurate predictions over time [25].

A more computationally feasible alternative to the complex Doyle-Fuller-Newman model is the single-particle model (SPM). SPM provides a simplified approach to modeling a lithium-ion cell's behavior by significantly reducing

the complexity of internal dynamics. Rather than accounting for detailed ion concentration profiles throughout the electrodes and electrolyte, the SPM assumes uniform lithium concentration within each electrode. This assumption simplifies the governing equations, as it reduces the partial differential equations in the DFN model to simpler ordinary differential equations, thereby enabling faster computations [61].

However, this simplification comes with limitations. A key limitation of the Single Particle Model (SPM) is its assumption of uniform concentration gradients within each electrode. This simplification can result in inaccuracies when predicting battery behavior under specific conditions, especially during high charging or discharging currents (C-rate ¿ 1) [66]. The C-rate represents the rate at which a battery is charged or discharged relative to its maximum capacity. At higher current densities, significant concentration gradients arise due to the rapid transport of lithium ions, leading to non-uniform distribution of lithium within the electrode particles. Since the SPM does not model these gradients, it cannot accurately capture the effects of ion depletion or accumulation in different regions of the electrode, leading to significant state estimation errors.

Additionally, the SPM does not explicitly model the electrolyte concentration and potential distribution across the cell. This omission limits its ability to capture polarization effects associated with electrolyte depletion, which are particularly pronounced under high-power charging and discharging. In applications where precise control of electrolyte dynamics is necessary, such as fast charging or discharging at high currents, this limitation can lead to significant deviations from actual battery behavior, potentially impacting safety and battery longevity.

Also, the SPM's accuracy can be limited by its inability to account for thermal effects, as temperature plays a significant role in influencing lithium-ion

transport, reaction kinetics, and cell degradation. In most implementations, thermal dynamics are either neglected or incorporated as an external parameter, which may not be sufficient for applications involving wide temperature ranges or rapid temperature fluctuations. As a result, the SPM's simplified approach can limit its utility in applications where both thermal and electrochemical stability are crucial.

In summary, while physical models like the DFN offer highly accurate state estimations and precise insights into battery behavior, they present substantial challenges for real-time applications due to their computational intensity. Although simpler models like the SPM improve computational efficiency, they may compromise accuracy under certain conditions. Therefore, for real-time applications, alternative modeling approaches that reduce complexity while maintaining sufficient accuracy across the battery's full operational range are generally preferred.

2.1.2 Direct Measurement Methods

The high modeling and computational costs associated with advanced SOC estimation methods have led to the adoption of simpler techniques, such as direct measurement methods. These techniques estimate SOC using measurable physical or chemical properties of the battery without relying heavily on computational models. These approaches are generally straightforward, such as voltage-based methods (§2.1.2) and current integration techniques (§2.1.2), but their accuracy and applicability can be limited under certain conditions.

Open Circuit Voltage Method

The open-circuit voltage (OCV) method [48–50], while straightforward and less computationally demanding, generally suffers from a limited operational range and sensitivity. The OCV method, for example, estimates SOC by using the battery's OCV, or Thevenin voltage in circuit terms, which is a critical parameter that reflects numerous aspects of the battery's internal state and overall performance. This method operates on the premise that the voltage at the battery terminals fluctuates with the SOC, reaching a maximum when the battery is fully charged and a minimum when it is fully discharged.

To implement this method, the OCV-SOC relationship must first be mapped, typically by subjecting the battery to a series of charge and discharge cycles under controlled conditions. These cycles allow for an empirical mapping that can then be stored as a static lookup table or approximated using polynomial equations [67]. However, this relationship is not fixed; it varies significantly with the battery's chemical composition, the rate of charging and discharging (known as the C-rate), and the ambient temperature. Consequently, a unique OCV-SOC mapping may be required for each set of operating conditions to achieve an accurate SOC estimate.

A particular challenge arises with lithium-ion batteries, which exhibit relatively flat charge/discharge curves across a broad SOC range. In these cases, the OCV changes very little across large portions of the SOC spectrum, making it difficult to pinpoint SOC from voltage readings alone. This limitation becomes even more pronounced in real-world applications, such as electric vehicles, where batteries are subject to rapidly changing operating conditions and dynamic current profiles. In contrast, this approach is more effective for other battery chemistries, such as lead-acid batteries, which exhibit relatively linear charge/discharge curves. This

linearity enables a more accurate estimation of the state of charge (SOC), as a measured voltage can be reliably correlated to a specific SOC value.

Additionally, standard OCV-SOC mappings are typically generated using static current levels during charging and discharging or with simple dynamic profiles that do not adequately stimulate the full range of the battery's internal processes. Such mappings, while useful in steady-state or controlled laboratory settings, often fall short of accurately representing the SOC in applications with complex and variable charge/discharge cycles. For electric vehicles, where high power demands and regenerative braking cycles lead to frequent fluctuations in current and load conditions, the SOC estimation based solely on OCV becomes increasingly unreliable. These discrepancies can result in suboptimal battery management, reducing both performance and lifespan.

As electric vehicle applications grow in scope and complexity, the limitations of traditional OCV-based SOC estimation methods highlight the need for more adaptable and robust modeling techniques that can dynamically adjust to the battery's operational conditions.

Coulomb Counting Method

The Coulomb counting method, also known as current integration, is a commonly applied approach for estimating the state of charge of lithium-ion batteries [51–53]. This method operates by measuring the current flow into and out of the battery over time, and integrating this current to estimate the net charge change. Normalizing the integrated charge by the cell's capacity yields SOC values within the range of 0% (fully discharged) to 100% (fully charged) [37]. This process makes Coulomb counting an intuitive approach since it directly links SOC to the net charge added or removed from the battery. As a result,

it can produce relatively accurate results under steady conditions, providing a continuous indication of SOC during battery operation.

The effectiveness of Coulomb counting, however, is highly dependent on the precision of current measurements, making it susceptible to drift and inaccuracies, particularly under real-world conditions. In complex and dynamic applications like electric vehicles, Coulomb counting's accuracy can degrade due to the inherent variability in operating conditions. Since Coulomb counting relies entirely on measured current values, any sensor drift or errors in the current measurement accumulate over time, leading to growing discrepancies in SOC estimation. For example, if there is even a slight deviation in current measurements over many cycles, the SOC estimate can become increasingly inaccurate, reflecting a phenomenon commonly referred to as drift.

Self-discharge is another significant challenge for Coulomb counting, as the method cannot directly account for charge losses when the battery is idle. Lithium-ion batteries are subject to gradual self-discharge over time, especially when left idle for extended periods. Since Coulomb counting lacks a feedback mechanism to adjust for this, it cannot correct for charge loss due to self-discharge, further compounding SOC inaccuracies.

In addition to drift, the SOC estimation accuracy of Coulomb counting is affected by the sampling rate of current measurements. A lower sampling rate may miss finer details in current fluctuations, reducing the method's accuracy, especially in dynamic charging and discharging environments like those in EVs. Since Coulomb counting is an open-loop method, it lacks an inherent feedback mechanism that could self-correct for cumulative errors. Consequently, any initial calibration errors continue to influence SOC estimation over time, which poses long-term accuracy challenges.

Coulomb counting is also impacted by variations in Coulombic efficiency, which is the ratio of charge discharged to charge charged during a full cycle [68]. Several factors, such as battery aging, ambient temperature, and discharge rates, affect Coulombic efficiency. For instance, as the battery ages, internal resistance increases, which can lead to losses during charge and discharge cycles, ultimately affecting Coulombic efficiency. Temperature fluctuations also impact the rate of chemical reactions within the battery, altering its effective efficiency during charging and discharging cycles [67]. Differences in charge and discharge rates (C-rates) can further compound these issues by introducing inconsistencies in Coulombic efficiency, as batteries may exhibit lower efficiency at higher current levels due to increased resistance and heat generation. Additionally, side reactions, such as electrolyte decomposition, can contribute to charge inefficiencies, leading to further inaccuracies in SOC estimates.

Due to these limitations, the accuracy of Coulomb counting often diminishes over prolonged use, especially in applications where precise SOC estimation is crucial. To counteract this, frequent recalibration of the current measurement system is required to minimize drift and error accumulation. However, implementing such a calibration regimen is costly and time-consuming, especially for high-demand applications.

2.1.3 Equivalent Circuit Models

Equivalent Circuit Models (ECMs) have been extensively studied in the field of lithium-ion battery modeling due to their balance of simplicity, computational efficiency, and adequacy in capturing basic battery behaviors [9, 15, 36, 37, 54, 55, 69–76]. ECMs are typically used for estimating crucial parameters such as state of charge, state of health, and voltage response, and they have been developed with

a range of complexities to meet different modeling needs. The basic framework of ECMs involves simplified electrical representations of battery behavior using standard circuit elements like resistors, capacitors, and voltage sources. These elements approximate the electrochemical processes within lithium-ion batteries, with parameters commonly identified through empirical testing [9, 36, 74]. For lithium-ion batteries, ECMs typically consist of a series resistance to represent the immediate voltage drop due to internal resistance and one or more RC (resistor-capacitor) networks to model dynamic responses, capturing delayed voltage changes from processes like ion diffusion and polarization. These models range from simple structures with a single RC pair to more complex configurations with multiple RC pairs, such as the second-order RC model shown in Fig. 2.2. More complex configurations can provide enhanced capability to simulate intricate

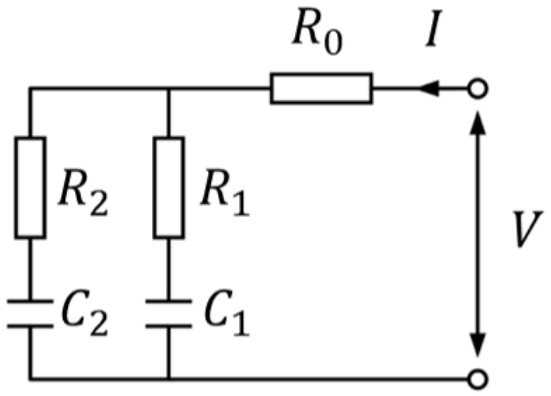

Figure 2.2. Diagram of Equivalent Circuit Model.

dynamic behaviors [77]. The applicability of ECMs can be categorized by their complexity. Simple ECMs with a single RC pair are typically adequate for low-dynamic applications, such as portable electronics, where capturing a general voltage profile is sufficient. In contrast, more complex ECMs that incorporate multiple RC pairs are better suited for high-power applications, as they can

capture transient behaviors and relaxation effects that occur after rapid current changes; essential for accurate modeling in electric vehicles (EVs) and other high-power contexts [24, 37, 76, 77]. In general, ECMs' computational efficiency, simple structure, and ease of parameterization make them widely popular in real-time battery management system (BMS) applications. Their computationally efficiency, makes them particularly appealing for embedded systems where processing power and memory are limited. For simple ECMs, modeling costs are low because parameters, such as resistance and capacitance, are easily identifiable through pulse discharge tests or similar experiments. This ease of parameterization makes ECMs suitable for various battery types and applications [9].

The adaptability of ECMs has enabled their widespread use in state of charge (SOC) and state of health (SOH) estimation [15, 36, 71, 77]. By adjusting basic parameters, ECMs can deliver sufficiently accurate results in scenarios that do not require detailed internal analysis. Moreover, with suitable tuning, ECMs can offer reliable predictions in short-term scenarios, where the battery operates under steady or predictable conditions. Despite their advantages, ECMs exhibit several limitations, particularly in applications that demand high accuracy across varying conditions or long-term state tracking. These limitations are widely recognized in the literature and relate primarily to their limited operational range, lack of degradation modeling, and strong sensitivity to environmental factors [10, 72].

ECMs often underperform outside the conditions for which they were calibrated, such as high C-rates or extreme temperatures. When encountering unforeseen conditions, ECMs may require recalibration of model parameters. For example, in [9], an ECM designed for a 21700 NCM811 lithium-ion cell was parameterized within an SOC range of 80% to 20% and specific C-rates. This constrained range limits their applicability in scenarios with variable operating

conditions, such as high-current discharges during EV acceleration, where complex electrochemical dynamics are prevalent. Moreover, ECMs can struggle to account for highly nonlinear behaviors under extreme conditions. During rapid acceleration or high-power discharge in EVs, for instance, ECMs with a single RC pair may fail to capture voltage sags accurately, leading to estimation errors [66, 77].

A significant drawback of ECMs is their inability to account for the effects of degradation mechanisms, such as lithium plating, electrolyte decomposition, or solid electrolyte interphase (SEI) growth. These limitations prevent ECMs from accurately tracking SOC and SOH over the battery's lifespan, which is essential for long-term battery management applications [66]. Unlike physics-based models that reflect changes in internal cell processes, ECMs provide overly simplified predictions, that can result in unreliable results in aged batteries where internal resistances and capacities change over time.

ECMs are also highly sensitive to temperature fluctuations, often requiring external compensation methods to maintain accuracy. For example, ECMs may fail to predict voltage accurately at low temperatures due to substantial increases in internal resistance. In cold-climate applications, extensive recalibration may be necessary to maintain prediction accuracy [66, 77]. Another challenge is parameter drift in ECMs over time as the battery undergoes cycling and aging. Frequent recalibration becomes necessary in high-cycle applications, such as grid storage and EVs, complicating maintenance efforts. Without adjustments, ECM predictions become less reliable as the battery ages, leading to potential estimation errors [10]. To mitigate parameter drift, continual recalibration is required, which is a burdensome and costly task, especially, for complex ECMs that incorporate multiple RC pairs.

In summary, ECMs offer a practical and computationally efficient approach for modeling lithium-ion battery behavior, making them valuable tools in real-time BMS applications where computational simplicity and short-term accuracy are sufficient. Their ease of parameterization and adaptability contribute to their widespread use in SOC estimation in moderate-power applications. However, ECMs are limited in scenarios with variable operating conditions, such as high C-rates or low temperatures, and are generally unsuitable for long-term degradation tracking due to their simplified structure and lack of connection to internal electrochemical processes. For applications requiring high accuracy, such as EVs with significant power demands, research increasingly favors hybrid or physics-based models that better capture the complex dynamics of lithium-ion batteries and extend their operational range. Continued research is essential to enhance ECMs or develop hybrid models that balance computational efficiency with long-term accuracy, especially for critical applications in energy storage and electric vehicles.

2.1.4 Machine Learning Models

Machine learning (ML) has become a transformative tool in the study and management of lithium-ion batteries, providing powerful, data-driven methods to predict battery behaviors, optimize performance, and extend lifespan. Leveraging large datasets, ML models can capture and model complex battery dynamics, such as state of charge (SOC), state of health (SOH), and degradation processes—factors critical to applications where reliability is paramount, including electric vehicles and grid storage [58, 66].

These ML models, often termed "black box" models [78, 79], develop mathematical representations of battery systems directly from empirical data, bypassing

the need for traditional, physics-based approaches. This data-centric approach shows substantial promise in accurately predicting battery behavior, making ML highly suitable for real-time battery management systems, where efficient processing is essential. Moreover, ML is not a one-size-fits-all solution but offers a flexible suite of tools adaptable to various stages of battery development, from classification and dynamic modeling to managing unforeseen operational scenarios through advanced techniques like ensemble learning and transfer learning [80].

Various machine-learning techniques have been explored to address the complex problem of battery performance management. These methods capitalize on abundant measurement data and include neural network (NN) frameworks, such as feed-forward NNs [56, 81, 82], recurrent NNs [57, 83], Elman NNs [84], stochastic fuzzy NNs [85], convolutional NNs [86, 87], backpropagation NNs [88], and nonlinear autoregressive NNs [89]. Additionally, support vector machine (SVM) frameworks have shown good performance in predicting SOC and SOH, including implementations that employ a moving window [90, 91] and hybrid approaches with fuzzy clustering techniques [92, 93]. Such ML-based models outperform many traditional approaches in their predictive accuracy for SOC and SOH.

For instance, in [94] a two-layer neural network (NN) with 30 neurons in the hidden layer could predict battery state with an error rate of around 4%, leveraging the NN's capability to model nonlinear battery behavior and capture complex interactions between voltage and other parameters. Similarly, Meng et al. [91] achieved comparable performance with an SVM strategy incorporating a moving window, which improved computational efficiency while modeling the battery.

However, standalone ML models often encounter limitations in accurately estimating SOC, especially under variable operating conditions. These challenges stem from their lack of integration with the underlying physics governing lithium-ion battery processes, insufficient data, and the inherent difficulty of ML models in handling out-of-distribution scenarios effectively. To increase accuracy and reduce RMS error, ML methods are frequently combined with traditional techniques or inference mechanisms, such as Kalman filters, to create hybrid models. Section §2.1.5 presents examples of these hybridization methods. For instance, combining a neural network model with an Extended Kalman Filter (EKF) allows the EKF to dynamically refine the neural network's SOC predictions by probabilistically merging them with real-time battery measurements, reducing the RMS error from 4% to 2%. This integration leverages the EKF's real-time updating capabilities, enabling the hybrid NN-EKF model to manage transient behaviors and nonlinearities in real-world battery operation, such as variable loads and temperature fluctuations. However, hybridization also introduces modeling complexity and can inherit limitations from the traditional methods it combines.

Alternatively, purely data-driven solutions, such as deep learning, offer another approach. Deep neural networks (DNNs), which involve multi-layered neural architectures, are particularly effective for handling large volumes of complex battery data. Deep Feedforward Neural Networks (DNNs), for example, can estimate SOC by directly mapping measurement inputs to SOC values. Training data, generated under a range of operating conditions in controlled laboratory settings, allows DNNs to learn complex dependencies within the data, encoding these patterns into network weights and yielding accurate SOC predictions. As a result, DNNs effectively capture the nonlinear behaviors of batteries under diverse conditions [56].

ML models can also analyze sensor data in real time, allowing dynamic adjustments to charging protocols, which is valuable in fast-charging applications where algorithms must balance speed and battery degradation. Despite these benefits, several limitations of ML in battery applications persist. High-quality and extensive datasets are typically needed to build reliable models, which poses challenges when data is scarce [24]. Limited data, for example, can lead to inaccuracies in state predictions.

Interpretability is another critical issue. Many ML models, especially deep neural networks, operate as "black boxes," making it difficult to understand the relationship between inputs and predictions [78]. This lack of transparency can be problematic for safety-critical applications, where understanding model outputs is essential for trustworthiness and regulatory compliance [58].

Moreover, computational complexity can also challenge ML's practical application in battery management. Training and deploying large ML models are resource-intensive and may require advanced computational hardware, which can be impractical in embedded battery management systems [76].

Another limitation is ML models' struggle to generalize across diverse operational conditions, where performance often degrades under new or variable environments not represented in the training data. In battery management, variations in temperature, charge/discharge rates, and battery aging introduce data distributions that may fall outside the training scope of machine learning models, potentially resulting in inaccuracies. This issue, known as the generalization gap, is particularly critical for applications such as electric vehicles and grid storage, where batteries operate under dynamic conditions, and reliable predictions are vital for safety and efficiency [95, 96]. While approaches such as transfer learning, domain adaptation, and ensemble methods can enhance

generalization, they increase system complexity and are not complete solutions. Addressing this challenge requires large, diverse training datasets that cover a wide range of operating conditions, a process that is often resource-intensive and involves significant costs and time investment.

In conclusion, while ML offers substantial potential for improving lithium-ion battery performance, data requirements, interpretability, computational demands, and generalizability remain active areas of research. Future studies will likely focus on refining hybrid models and extending ML's applicability to different battery chemistries and operating environments, advancing the reliability and efficiency of ML-driven battery management systems.

2.1.5 Hybrid Models

Hybrid battery modeling approaches are gaining traction as they combine physics-based, equivalent circuit, and data-driven models to improve the performance, robustness, and complexity management of battery systems. Rather than treating these methodologies as mutually exclusive, researchers are increasingly exploring integrated or "gray box" models, particularly those that blend physics-based and data-driven techniques. [97] provides a comprehensive overview of methods that unite physics-based models with machine learning, illustrating how hybrid models can combine the accuracy and efficiency of data-driven insights with the reliability of established physical principles. These models capture complex battery behaviors while grounding predictions in physical laws, enhancing both interpretability and adaptability [98].

In contrast to purely data-driven "black box" models, gray box models use physics-based principles to retain transparency and theoretical grounding, offering a more interpretable framework. For instance, physics-informed neural networks

(PINNs) enforce adherence to physical laws like thermodynamics and charge conservation, helping the model remain consistent with established scientific knowledge [99]. Similarly, physics-based feature engineering can embed domain knowledge into model inputs, allowing data-driven methods to better learn and predict meaningful patterns [81]. Another approach is physics-based model correction, which uses data-driven techniques to refine traditional physics-based models, improving accuracy in capturing complex or partially understood phenomena [81]. Additionally, physics-based activation functions in neural networks incorporate known physical relationships directly into the learning process, enhancing model fidelity [66].

These hybrid approaches are essential for advancing battery technology, as they provide more reliable, interpretable, and generalizable models suited for real-world applications.

Several hybrid SOC estimation methods also integrate filtering algorithms like Extended Kalman Filter (EKF) [24, 71, 100, 101], Unscented Kalman Filter (UKF) [69, 102, 103], Luenberger observer [104], sliding mode observer [105, 106], and Adaptive observer [76, 107, 108] to address measurement and modeling uncertainties [70, 76]. These filters are particularly valuable in managing noise and drift associated with direct measurement methods, such as open-circuit voltage (OCV) and Coulomb counting, which tend to experience cumulative errors without ground truth data. Filter-based hybrid models can offer robust prediction performance, although they can be sensitive to numerical errors over time, impacting variance parameters and leading to filter divergence. Rigorous model development and noise characterization can mitigate these issues but may increase computational costs.

In response to limitations with OCV and Coulomb counting, various hybrid methods integrate these with state-based models combining electrochemical or equivalent circuit models with measurement and machine learning methods. Some models use OCV resets to mitigate SOC estimation drift, leveraging open-circuit voltage to recalibrate SOC predictions and manage errors due to initial SOC uncertainties [109–111]. While these hybrid models improve SOC prediction, they can be sensitive to operating conditions like temperature, restricting the battery's operational range.

In conclusion, hybrid models offer an efficient solution for battery modeling, enhancing predictive accuracy while addressing specific limitations of purely data-driven or physics-based approaches. However, further advancements in hybrid methods are needed to extend their accuracy across broader operational conditions and to reduce computational costs, ultimately improving their applicability in battery management systems across diverse applications.

2.1.6 Explicit Data-driven Models

The limitations in existing methodologies for modeling lithium-ion batteries have motivated our research toward an alternative, data-driven approach. Data-driven modeling (DDM) of complex systems, like LiBs, offers substantial promise for accurately capturing the underlying dynamics by leveraging explicit terms derived directly from input-output data [10]. This approach enables enhanced prediction and understanding of system behavior [112–114]. DDM aims to address the challenges of traditional methods by using accessible data to uncover interpretable models that are well-suited for system identification, prediction, and control. Moreover, recent advances in machine learning further support this approach by enabling the creation of interpretable models that bridge data and

physical insights, facilitating efficient control design and enabling the desired performance of complex systems [112].

The field of data-driven modeling has been approached through various frameworks and for a variety of applications including turbulence, epidemiology, neuroscience, and finance, where systems are high-dimensional and nonlinear and exhibit multiscale phenomena in both space and time [112, 115]. Since the 1960s, various data-driven modeling frameworks have been developed to capture dominant system characteristics by leveraging input-output data and state-space models, beginning with Kalman's introduction of minimum realization (Kalman decomposition). This seminal method proposed the realization of state space models of linear systems via analysis of experimentally obtained response data and established the important principles of realization theory in terms of system controllability and observability [116–118]. Following Kalman's minimum realization, many model reduction techniques were introduced for the development of stable reduced-order models (ROM) of the full high-dimensional systems. The fields of data-driven modeling and model reduction are intertwined, with methods of one field often using the developments of the other for inspiration.

Balanced Truncation (BT) is a well-established model-reduction technique in the field of control theory for relatively small linear input-output dynamical systems [119–121]. BT performs a balancing coordinate transformation using the concepts of minimum realization to balance the observable and controllable components of the system, yielding tractable ROM, which captures the important aspects of the full-order dynamics. However, BT is expensive and potentially intractable for high-order systems since it requires the computation of the system's controllability and observability Gramians [119].

An efficient alternative is proper orthogonal decomposition (POD), which is a model reduction technique developed for high dimensional systems [122]. POD works by performing a coordinate transformation to an orthogonal basis and offers improved computational efficiency but may result in unstable models even for stable systems [122–124].

The emergence of efficient computational methods for Singular Value Decomposition (SVD) [125–127] inspired new data-driven modeling approaches to the realization problem, such as the Eigensystem Realization Algorithm (ERA). It was developed for modal parameter identification of aerospace structures such as the Galileo spacecraft from measured data [128]. ERA extends the concepts of minimum realization in combination with the SVD technique for treating noisy data and model reduction. It provides accurate low-order linear models of high-dimensional systems under the restrictions that the system operates in a linear range, has time-invariant dynamics, and allows for computation of its impulse response [116, 118, 128, 129].

The Observer Kalman Identification (OKID) technique was developed as an extension to ERA that lifts some of its restrictions by using an asymptotically stable observer to estimate the system's impulse response from any set of pseudo-random inputs [130–132]. ERA and OKID are suitable for input-output systems with higher rank (dimensionality) than the number of observables [131, 133]. After OKID, the tractable model reduction technique, balanced POD (BPOD), was introduced. BPOD combines the balancing principle of BT with the computational efficiency of POD. It performs the balanced truncation through direct and adjoint impulse response functions, which efficiently approximate the system's Controllability and Observability Gramians. However, BPOD is limited to systems with

known models due to the required computation of adjoint response data, which can only be obtained through manipulation of existing models [122, 134, 135].

An adaptive learning alternative is Jacobian Learning, which utilizes learning methods to identify and recursively update a system's input/output sensitivity (dominant characteristics) from measurement data [136, 137]. The Jacobian learning process is carried out via a recursive least squares approach [112]. Once the sensitivity is learned, this approach allows for model-free control of the system without explicit knowledge of the underlying dynamics [138]. However, to maintain the accuracy of the learned Jacobian, it must be recursively updated, and thus it is best suited for offline applications.

Dynamic Mode Decomposition (DMD) is a computationally efficient data-driven modeling technique for identifying linear reduced-order models of complex high-dimensional systems. It produces coupled sets of spatial-temporal modes (structures) that dominate the observed measurement data. These structures are connected by a linear dynamical system that demonstrates their evolution in time, and they are identified by approximating the system's leading eigen-decomposition [133, 139–144]. Recent extensions such as eDMD (Extended Dynamic Mode Decomposition) use the DMD technique as the computational machinery to approximate the Koopman operators of highly nonlinear systems [145–148].

Koopman operators characterize the nonlinearities of a system via a transformation to an intrinsic coordinate system where nonlinear dynamics appear linear [149–153]. Obtaining linear representations of strongly nonlinear systems via approximation of Koopman operators has the potential to revolutionize our ability to predict and control complex systems [153]. However, this often requires manual preparation of nonlinear observables according to the system's underlying physics, which is often unknown. Furthermore, automated learning approaches

for Koopman operators [152] are often only suitable for purely predictive models that do not account for external forcing (control inputs).

Nonlinear data-driven modeling techniques such as SINDy (Sparse Identification of Nonlinear Dynamics) [154] provide an efficient alternative to the previously detailed linear modeling approaches (e.g. ERA, OKID, DMD, etc.) that allow for accurate characterization of nonlinear dynamics and inclusion of the effects of external forcing to the identification problem via extensions such as SINDYc (SINDY with control) [155]. SINDy models the governing equations of a dynamical system by constructing a library of candidate terms—linear and nonlinear transformations of the measurement data—and assigning coefficients that indicate the significance of each term in explaining the system's behavior [154–160]. This approach uses a sparsity-promoting algorithm to optimize the model weights, selecting a parsimonious subset of basis functions that accurately represents the data with minimal terms [113, 161].

SINDy's capability to create computationally efficient (sparse) models of complex, high-dimensional, and nonlinear dynamical systems, while capturing multiscale phenomena across both space and time, makes it particularly well-suited for modeling lithium-ion battery behavior. Its control-oriented framework, which accounts for external forcing, produces interpretable models that accurately capture the underlying dynamics, offering valuable insights into system behavior. These attributes make SINDy an appealing and practical tool for real-time applications in battery management systems . Furthermore, this explicit data-driven modeling approach bypasses the need for proprietary or inaccessible data from mechanistic models, delivering accurate and practical battery representations. It also demands significantly less data compared to machine learning methods like

neural networks, while producing interpretable battery digital twins well-suited for control-oriented applications.

Despite its strengths, SINDy encounters notable challenges when applied to the state of charge (SOC) dynamics of Li-ion batteries, particularly in maintaining accuracy across diverse operating conditions, including varying ambient temperatures. Battery dynamics, such as temperature-dependent changes in internal resistance, are inherently complex. Moreover, the slow evolution of these dynamics limits effective system excitation across low and high temperatures within a single experimental dataset [77].

Several limitations of SINDy arise from its reliance on generic libraries (e.g., polynomial terms), which may be adequate for simple or well-defined problems but struggle with more complex systems. These libraries can yield inaccurate representations of the data, and the choice of sparsification parameters (hyperparameters) often results in significantly different models. Furthermore, the method may misfit the data to a nonlinear model, even when using an appropriate library, as identifying the "correct" set of nonlinear functions is inherently challenging. Different combinations of terms can produce similar input-output behaviors, complicating the modeling process.

Identifying an optimal model that balances accuracy and complexity requires solving an outer-loop optimization problem involving the hyperparameters of the sparsification algorithm. The effectiveness of the learning algorithm depends heavily on the selection of these hyperparameters, which govern the learning process and influence the resulting model. Proper hyperparameter selection is crucial for achieving optimal performance, but the vast search space often leads to suboptimal outcomes. Hyperparameter tuning, therefore, necessitates an iterative

outer-loop optimization process to fine-tune these parameters [162], as their impact on model quality is inherently non-deterministic.

Thus, the generic methodology must be refined to optimize the modeling of SOC dynamics. A significant advantage of this approach, however, lies in its high adaptability, enabling specialization to meet the unique requirements of battery systems.

In recent work, I extended the SINDy modeling techniques to energy storage systems, with a particular focus on lithium-ion batteries [10, 13, 14, 66, 66, 163].

This work introduces the Physics-Informed and Temperature-Dependent Explicit Data-Driven (PhITEDD) framework, a digital twin for Li-ion battery state-of-charge (SOC) dynamics designed for real-time monitoring and control. By integrating concepts of mechanistic and data-driven approaches, the framework achieves a balance between accuracy, computational efficiency, and interpretability. Key innovations include incorporating physics-informed terms to reduce error under unseen conditions, a Monte Carlo library search to identify nonlinear terms, and automated hyperparameter tuning for optimal model complexity and accuracy. Recalibration strategies ensure consistent performance across diverse operating conditions, including wide temperature ranges. PhITEDD offers a robust, efficient, and interpretable solution for SOC estimation, capturing nonlinear dynamics and temperature dependencies directly from measured data, making it ideal for advanced battery management systems.

Building on the comparative study conducted in [58], the PHITEDD framework is evaluated against the most promising battery SOC modeling and prediction methods highlighted in the study. Fig. 2.3 showcases its advantages across key performance criteria, including:: (i) Interpretability: the comprehensibility of a model's decision-making process. (ii) Robustness: a model's ability to accurately

Term	Definition
BMS	Battery Management System
BPOD	Balanced Proper Orthogonal Decomposition
BT	Balanced Truncation
C-rate	rate at which a battery is fully charged or discharged
DDM	Data-Driven Modeling/Model
DFN	Doyle-Fuller Newman, physics-based FOM of LiB
DMD	Dynamic Mode Decomposition
ECM	Equivalent Circuit Model, the electrical model of LiB
ERA	Eigensystem Realization Algorithm
ESS	Energy Storage System
EV	Electric Vehicle
JL	Jacobian Learning
KOT	Koopman Operator Theory
LiB	Lithium-ion Battery
ML	Machine Learning
NN / DNN	Neural Network / Deep Neural Network
OCV	Open Circuit Voltage
OKID	Observer Kalman Identification
POD	Proper Orthogonal Decomposition
RMSE	Root Mean Square Error
ROM	Reduced-Ordered Model
SOC	State of Charge [%]
SPM	Single-Particle Model, physics-based ROM of LiB
SVD	Singular Value Decomposition
SINDY	**Sparse Identification of Nonlinear Dynamics**

Table 2.1. Battery and Modeling Acronyms

simulate battery behaviors beyond the scenarios represented in the training data. (iii) Fidelity: the precision with which a model replicates the underlying physics that dictate battery behavior. (iv) Transferability: a model's capability to apply across multiple battery chemistries without needing significant modification. (v) Computational efficiency: the time or memory needed to perform a calculation. (vi) Data efficiency: the extent to which a battery model can be parameterized using a minimal amount of measured data. (vii) Domain Knowledge: the essential understanding of the underlying principles and physics that govern the system's behavior. (viii) Applicability: the range within which a model can effectively operate, including the data space it can accurately handle.

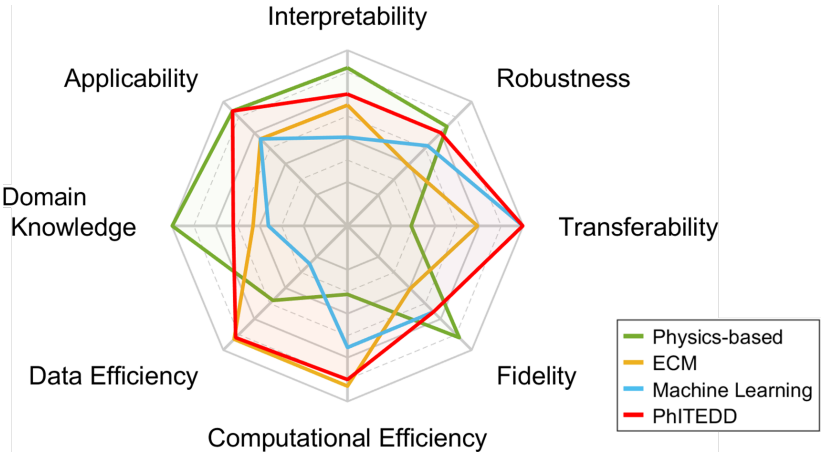

Figure 2.3. Comparison of SOC Modeling and Prediction Methods.

2.2 Fast Charging

Over the last decade, LiBs have become the technology of choice for grid storage, portable electronics, and specifically electric vehicles (EVs). However, despite advancements in battery technology and incentives like tax credits, EV adoption still faces a major hurdle in slow charging times. Charging an EV battery

pack to full capacity takes significantly longer than refueling a conventional vehicle [23]. This has led to increased demand for enhanced battery technologies that deliver fast-charging protocols with minimal charging duration while ensuring safety during operation.

The cycle life of lithium-ion batteries is significantly affected by the selected charging protocol [164]. Furthermore, fast charging can accelerate battery degradation. Thus, a trade-off exists between charging speed and battery lifespan [165]. The primary risk comes from subjecting the battery to high C-rates and the elevated temperatures that a fast charge generates [166]. Elevated temperatures can accelerate electrochemical aging, resulting in increased lithium plating, higher mechanical stresses, and an increased rate-of-growth of the SEI (solid-electrolyte interface) layer [18, 21]. The fast charging problem has been explored through various methods, that can be grouped into two categories passive charging strategies (model-free) detailed in §6.1 and active charging strategies (model-based) detailed in §2.2.2.

2.2.1 Passive Charging

Passive charging strategies, including constant-current (CC), constant-current constant-voltage (CC-CV) [166–168], multi-stage CC-CV [169], and pulse charging techniques [21], represent traditional approaches for charging lithium-ion batteries. These methods follow fixed charging profiles defined by constraints on current, voltage, or power, and do not adapt to the dynamic state of the battery during the charging process. This model-free nature makes them straightforward to implement but inherently heuristic [23].

For example, the CC strategy applies a constant current throughout the charging process, while the CC-CV approach transitions to a constant voltage phase

once the battery reaches a specified threshold. Multi-stage CC-CV extends this concept by introducing additional stages to optimize the charging time and reduce stress on the battery. Pulse charging, on the other hand, alternates between charging and rest phases to mitigate heat buildup and enhance lithium-ion diffusion within the electrode. Despite their simplicity and reliability, these methods often overlook the battery's internal electrochemical dynamics and thermal responses, which can lead to suboptimal charging efficiency and accelerated degradation.

These limitations have spurred the development of active optimal charging protocols, which adaptively adjust charging profiles based on real-time feedback from the battery.

2.2.2 Active Charging

Active protocols aim to balance fast charging demands with the need to minimize adverse effects on battery health, such as capacity fade and increased internal resistance. By leveraging insights from battery modeling and control, active strategies offer a more tailored approach, considering factors such as state of charge (SOC), temperature, and internal resistance. This shift toward active methods represents a critical advancement in achieving efficient, safe, and durable fast-charging solutions.

These charging protocols can be split into two categories. The first category uses empirical battery models such as equivalent circuit models (ECMs) [54] or machine learning models [56] to predict battery states using past measured data and state observers such as Kalman filters [170], or moving horizon estimators [171] to estimate the true/internal battery states.

Also, it includes a control or optimization scheme, such as linear quadratic control [172], Pontryagin's minimum principle [173], or model predictive control

(MPC) [174,175], to improve charging performance. A significant body of literature employs MPC to address the optimal charging problem. This problem is framed as a constraint-based optimization whose goal is to either minimize the time required to reach a specific state of charge (SOC) or maximize the SOC achieved within a set charging duration. However, this technique is known for its computational intensity. Furthermore, real-time implementations often use empirical models, which are unable to reflect physics-based parameters and compromise the physical precision of the solution [174].

The second category of optimal charging strategies involves using physics-based models for calculating the battery states. These methods often utilize an MPC control scheme along with reduced-order methods such as the single-particle model (SPM) [176], and electrochemical models with a constant electrolyte concentration [177] since they experience reduced computational complexity when compared to full-order alternatives. This approach allows for formulating a closed-loop optimization problem to minimize charging time and can more naturally include physics-based constraints. Nonetheless, its efficacy is hampered by model inaccuracies stemming from the simplified representation of the battery dynamics, alongside its considerable computational complexity. Moreover, the streamlined dynamics fail to exploit the system's capabilities, potentially resulting in a conservative or infeasible solution depending on the problem formulation.

This study proposes an adaptive learning and control strategy [112, 178, 179] to optimize battery charging profiles for minimal charge time while maintaining safety constraints on maximum cell temperature and voltage to prevent over-charging and overheating the battery, respectively. Using data from the full-order electrochemical Doyle-Fuller-Newman (DFN) model and a thermal energy balance model, the approach closely aligns with actual battery behavior [10, 28]. A hybrid

control framework maximizes initial current and dynamically transitions operating modes to meet constraints. The method begins with a known solution, such as CC-CV, and optimizes control points [180] to achieve fast, safe, and efficient charging.

2.3 Research Objectives and Approach

The primary research objective of this study is to address critical challenges in battery management systems by enhancing their performance, applicability, and adaptability under extreme temperatures, high C-rates, and varying states of charge, which are key factors influencing battery modeling, control, and overall performance. This includes developing accurate and computationally efficient models for characterizing state-of-charge (SOC) dynamics of lithium-ion batteries (LiBs) and devising optimal fast-charging strategies suitable for real-time applications, such as electric vehicles.

The electric vehicle (EV) market has experienced steady growth over the past decade, fueled by advancements in battery technology, stricter emissions regulations to combat environmental pollution, and increasing consumer demand for cleaner and more sustainable transportation. This shift towards electrification is further supported by heightened awareness of climate change and the environmental drawbacks of internal combustion engine (ICE) vehicles, including greenhouse gas emissions and air quality degradation. Consequently, EV adoption is anticipated to continue its upward trajectory.

As EV adoption grows, the demand for optimized energy storage systems, particularly lithium-ion batteries, is increasing. While lithium-ion batteries offer high energy density, low self-discharge rates, and decreasing costs, their performance is affected by factors such as temperature fluctuations, charge/discharge cycles, and

aging, which can degrade capacity over time. Optimizing performance, enhancing operational safety, and extending the longevity of electrified systems rely on the continued development and refinement of battery management systems (BMS). As a critical technology, BMS is essential for monitoring battery health, ensuring safety, and maximizing efficiency through optimized charging and discharging processes. To achieve optimal performance, BMS must have insight into the internal battery state, particularly the SOC, which functions as the equivalent of a fuel gauge in conventional vehicles. However, because direct SOC measurements are not possible, it must be estimated using measurable signals such as electrical current (I). Furthermore, optimizing charging performance presents new challenges and demands for BMS. Fast charging technology significantly reduces recharge time, enhancing convenience for long-distance travel and improving the user experience. However, the increased speed of charging places substantial pressure on BMS to carefully balance factors such as temperature regulation, charge rate optimization, and safety protocols, all while preserving battery longevity. To meet these heightened demands, BMS must implement highly adaptive charging strategies that effectively mitigate risks such as excessive heat generation, which can lead to thermal runaway and compromise battery safety.

The challenge of SOC estimation has been approached through various modeling techniques, including full-order electrochemical models and reduced-order models (ROMs). Electrochemical models, based on first principles, offer high-fidelity state estimates but require detailed knowledge of battery composition and are computationally intensive, making them unsuitable for real-time applications. In contrast, ROMs provide simplified representations of specific dynamic processes (e.g., SOC) using methods such as physics-based or data-driven approaches. While more efficient, ROMs often sacrifice some accuracy for computational simplicity.

This work presents a data-driven framework for efficient and accurate SOC estimation by developing a reduced-order nonlinear model that characterizes the complex dynamics of lithium-ion batteries using non-invasive input/output data. The approach extends the operational range of LiBs by incorporating SOC- and temperature-dependent behaviors, ensuring reliable performance across diverse conditions, including the highly nonlinear regimes of low temperatures and low SOC.

We accomplish this by leveraging state-of-the-art data-driven modeling techniques, such as SINDyC, augmented with tools from machine learning, control theory, and physics-informed learning. These advancements enable precise characterization of LiB nonlinear dynamics using operando input/output data (real-time measurements under working conditions) for training and validation. Our methodology applies a sparsity-promoting algorithm to approximate governing equations through a library of candidate terms. This library includes linear and nonlinear features of the measurement data, with the coefficients of these terms indicating their relative importance in capturing the underlying battery dynamics.

The optimal charging problem has been explored through various methods, broadly categorized into passive (model-free) and active (model-based) charging strategies. Since the cycle life of lithium-ion batteries is heavily influenced by the chosen charging protocol, achieving an optimal balance between charging speed and battery longevity is crucial. Fast charging, while enhancing convenience, can exacerbate battery degradation. The primary risks stem from exposing the battery to high C-rates and the elevated temperatures generated during fast charging. These elevated temperatures accelerate electrochemical aging, leading to reduced capacity and diminished overall performance over time.

Passive charging strategies, such as the constant-current constant-voltage (CC-CV) method, are traditional approaches for charging lithium-ion batteries that rely on fixed charging profiles constrained by current, voltage, or power limits. CC-CV, the most widely used technique, is favored for its model-free design, which enables straightforward implementation. This method delivers a constant current during the initial charging phase and transitions to a constant voltage once a predefined threshold is reached. However, because passive strategies do not adapt to the battery's dynamic state, they remain inherently heuristic and often yield suboptimal performance, particularly under varying operating conditions.

Active charging strategies provide a more tailored approach by incorporating factors such as SOC, voltage, and temperature into the charging process. These methods leverage battery models in combination with control or optimization schemes to enhance charging performance. The fast-charging problem is typically formulated as a constraint-based optimization, aiming to minimize charging time while mitigating adverse effects on battery health. However, the complex, nonlinear, and temperature-dependent dynamics of batteries pose significant challenges for active strategies. Model Predictive Control (MPC) is a widely used technique due to its ability to manage complex systems. Nonetheless, its computational complexity often necessitates the use of overly simplified battery models. Such models may lack the ability to accurately represent physics-based parameters, leading to reduced accuracy and limited operational ranges (e.g., C-rates, temperatures, and SOC levels). Consequently, these simplifications can compromise the physical realism of the solution, or in some cases, produce conservative or infeasible results depending on the problem formulation.

This work proposes an adaptive learning and control strategy to optimize the charging profile (electrical current) for minimizing battery charging time while

adhering to constraints such as maximum cell temperature and voltage limits. The proposed direct data-driven control framework dynamically adjusts charging profiles by integrating advanced learning algorithms with optimization techniques, leveraging feedback from the battery's state response. The framework operates by learning the Jacobian of the closed-loop system directly from input/output data, enabling the optimization of battery response based on its dynamic behavior. This approach provides an efficient and adaptive solution tailored to the battery's operational conditions, allowing precise control of the charging current at every stage of the process. By dynamically adjusting the charging profile, the framework reduces thermal stress commonly associated with fast charging, mitigating risks like overheating and thermal runaway. It also prevents degradation mechanisms such as lithium plating, solid electrolyte interface growth, and increased internal resistance, thereby enhancing overall charging efficiency and preserving battery performance and lifespan. A key advantage of this data-driven control approach is its versatility, making it applicable to a range of modeling methodologies, including physics-based models, data-driven frameworks, and real-world battery data.

Thus, by integrating advanced SOC prediction with an optimized fast-charging method, this work addresses critical challenges limiting the widespread adoption of electric vehicle technology. It provides a comprehensive solution that enhances safety, efficiency, and battery longevity, paving the way for more reliable energy storage systems that support the future of electric mobility.

This work has the following general hypothesis:

Hypothesis 1:

The complex state of charge dynamics in lithium-ion batteries, influenced by

varying operating conditions (e.g., temperature, C-rate, cell chemistry), can be efficiently represented by a nonlinear model derived from measured input/output data (voltage, current, and temperature) using a sparsity-promoting algorithm to identify key terms from a library of potential candidates.

Hypothesis 2:

An optimized fast-charging strategy, guided by a data-driven control framework, can minimize charging time while adhering to safety and operational constraints, such as maximum cell temperature and voltage limits .By dynamically modulating the charging current in response to real-time feedback from the battery's state (e.g., SOC, temperature, and voltage), this approach improves thermal management and mitigates electrochemical aging, thereby preserving battery capacity and sustaining performance over time.

To test these hypothesis, we have the following aims::

Aim 1:

Investigate the accuracy and efficiency of available approaches used to tackle the SOC estimation problem.

Task 1.1: Investigate the physics that governs the dynamics of LiBs (full-order modeling via first principles).

Task 1.2: Investigate available efficient alternatives for SOC modeling (reduced-order modeling) and identify their drawbacks and limitations.

Task 1.3: Compare the accuracy and efficiency of different modeling approaches on simulated and experimental data.

Aim 2:

Develop the procedures for generating and collecting simulated and experimental LiB input/output data.

Task 2.1: Investigate LiB charging and discharging routines for fast applications such as electric vehicles (EVs).

Task 2.2: Conduct charge and discharge tests on a simulated battery at varying operating conditions (e.g., SOC levels and cell temperatures).

Task 2.3: Design of Experiments to Attain a Valid Model Across Operating Conditions (e.g., SOC levels and cell temperatures).

Aim 3:

Develop the procedures and methodology for identifying reduced-order SOC models from input/output data that captures the LiB's SOC and temperature-dependent performance.

Task 3.1: Investigate available data-driven modeling methods to identify the technique best suited for modeling LiB's SOC dynamics.

Task 3.2: Use the identified methodology to develop preliminary reduced-order SOC models on simulated battery data at standard operating conditions and identify the limitations of the generic approach to energy storage systems.

Task 3.3: Enhance the generic methodology with tools from machine learning and control theory for improved predictive performance and improved coverage of the LiB's operational range (e.g., wide temperature range); and introduce domain knowledge to the learning process by curating

a set of physics-inspired terms to better and more efficiently capture the SOC dynamics.

Task 3.4: Develop an accurate and efficient data-driven SOC model that captures the LiB's SOC and temperature-dependent performance using our enhanced methodology and experimental battery data; and test the generalizability (real-world predictive performance) of the model on unseen experimental data corresponding to a wide range of operating conditions (e.g., different temperatures)

Aim 4:

Investigate the effectiveness, applicability, and limitations of current methods for solving the battery charging optimization problem.

Task 4.1: Explore the physical and electrochemical processes involved in charging lithium-ion batteries at high C-rates and their impact on battery performance and health.

Task 4.2: Explore existing fast-charging approaches, encompassing both model-free and model-based methods.

Task 4.3: Assess and compare the performance and limitations of existing fast-charging methods.

Aim 5:

Develop an optimized fast-charging strategy for lithium-ion batteries.

Task 5.1: Perform battery charging tests at varying C-rates using different methodologies, such as CC-CV and pulse charging while leveraging the full-order dynamics of the physics-based DFN battery model.

Task 5.2: Develop optimization criteria and operational constraints informed by existing fast-charging literature and the specific characteristics of the battery chemistry.

Task 5.3: Optimize the charging profile to minimize charge time while mitigating adverse impacts on battery health.

3. METHODOLOGY

This chapter details the methodology for developing a battery digital twin to capture state-of-charge (SOC) dynamics §3.1. The digital twin utilizes a reduced-order model composed of a library of candidate terms, with coefficients representing the relative significance of each term in describing the system behavior. By leveraging operando input/output measurements, this data-driven approach identifies governing equations through a sparsity-promoting optimization algorithm.

Additionally, this chapter introduces the direct data-driven control strategy employed to infer the battery's Jacobian matrix from input/output data §3.2. This strategy enables the optimization of the charging current profile, minimizing charging time while adhering to critical safety constraints, such as maximum cell temperature and voltage, effectively addressing the fast-charging optimization challenge.

3.1 Data-Driven Battery Model

This section details the methodology used to develop the temperature-dependent battery digital twin for SOC dynamics. A digital twin serves as a virtual dynamic representation of a system, allowing real-time monitoring and control. Its goal is to establish a strong link between the physical and digital realms, offering a potential solution for integrating mechanistic and data-driven approaches [70, 181, 182].

Here, the digital twin consists of an efficient and accurate reduced-order model, built from a library of candidate terms and corresponding coefficients that indicate the relative importance of each term in describing the data. Leveraging

battery operando input/output measurements, this data-driven approach uncovers governing equations using a sparsity-promoting optimization algorithm. This approach minimizes the need for extensive knowledge of the physical and material properties of the battery while integrating domain expertise to simplify the model and improve predictive accuracy, significantly lowering the costs of model development and implementation.

This section is structured as follows: the explicit data-driven modeling method is described in §3.1.1, the hyperparameter autorunning method is presented in §3.1.2, the Monte Carlo (random) search approach for library terms is described in §3.1.3, the re-calibration method for optimizing model coefficients on new data encompassing distinct operating conditions is described in §3.1.4, and the battery digital twin is presented in §3.1.5.

3.1.1 Explicit Data-Driven Modeling

The problem formulation presented in this section illustrates the approach to discovering governing equations by sparsifying a library of candidate terms. The nomenclature for this section is summarized in Table 3.1.

Table 3.1. Data-driven Modeling Nomenclature

Symbol	Description
$\mathbf{X}$	data matrix of time snapshots of states/model-outputs $\mathbf{x} = [\mathbf{SOC}]$
$\mathbf{U}$	data matrix of time snapshots of model-inputs $\mathbf{u} = [\mathbf{I}, \mathbf{V}]$
$\Theta(\mathbf{X}, \mathbf{U})$	feature library of $\mathbf{X}$ and $\mathbf{U}$
Ξ	set of sparse coefficients/weights
$P(\Xi)$	sparsity promoting penalty
λ	regularization hyperparameter
ξ_{th}	threshold hyperparameter

Here, the dynamic nonlinear model is assumed to take the following form:

$$\mathbf{x}[\mathbf{k}+\mathbf{1}] = f(\mathbf{x}[\mathbf{k}], \mathbf{u}[\mathbf{k}]) \tag{3.1a}$$

$$\mathrm{SOC}[k+1] = f(\mathrm{SOC}[k], I[k], V[k]), \tag{3.1b}$$

where the states/outputs (e.g., SOC) and inputs (e.g., I, V) are represented by $\mathbf{x} \in \mathbb{R}^{\mathbf{p}}$ and $\mathbf{u} \in \mathbb{R}^{\mathbf{q}}$ respectively. The function $f(\mathbf{x}[\mathbf{k}], \mathbf{u}[\mathbf{k}])$ represents the governing dynamics, which is assumed to be characterized by a few active terms. These active terms are chosen from a feature library comprising linear and nonlinear transformations of dynamic data via a sparse regression technique. A detailed discussion regarding the selection of library terms for LiBs is presented in 5.1. The feature library is developed using input/output measurements of the battery system. The collected data is structured as matrices (3.2) and (3.3) consisting of m time snapshots of $\mathbf{x}$ and $\mathbf{u}$, respectively.

$$\mathbf{X} \overset{\text{def}}{=} \mathbf{X}[\mathbf{k}, \mathbf{m}] \overset{\text{def}}{=} [\mathbf{x}[\mathbf{k}], \ldots, \mathbf{x}[\mathbf{k}+\mathbf{m}-\mathbf{1}]]^{\mathbf{T}} \tag{3.2}$$

$$\mathbf{U} \overset{\text{def}}{=} \mathbf{U}[\mathbf{k}, \mathbf{m}] \overset{\text{def}}{=} [\mathbf{u}[\mathbf{k}], \ldots, \mathbf{u}[\mathbf{k}+\mathbf{m}-\mathbf{1}]]^{\mathbf{T}} \tag{3.3}$$

The explicit data-driven model can be formulated as follows:

$$\mathbf{X}' = \boldsymbol{\Theta}(\mathbf{X}, \mathbf{U})\boldsymbol{\Xi}, \tag{3.4}$$

where $\Theta \in \mathbb{R}^{m \times D}$ is the library of candidate terms, $\Xi \in \mathbb{R}^{D \times p}$ is a sparse vector of coefficients indicating the relative importance of each term in describing the data, D is the number of library terms, and $X' \overset{\text{def}}{=} X[k+1, m]$ is a shifted temporal matrix of $\mathbf{X}$.

The coefficients of (3.4), Ξ, are used to enforce sparsity in the model. Sparse modeling is desired as it helps achieve interpretable and generalizable models by balancing accuracy and complexity. Moreover, sparsity helps prevent overfitting by reducing the number of active terms, which in turn improves efficiency. The sparse optimization problem for identifying the set of model coefficients is given by:

$$\min_{\Xi}. \ \mathcal{L}(\Xi) \stackrel{\text{def}}{=} \left(\left(X' - \Theta(X,U)\Xi\right)^2 + \mathcal{P}(\Xi) \right), \tag{3.5}$$

where $\mathcal{P}(\Xi)$ is the cost to promote sparsity. This work employs the sequentially thresholded Ridge regression (STRidge) algorithm to optimize (3.5). This regularizer defines $\mathcal{P}(\Xi) = \lambda \left\|\Xi\right\|_2 + \xi_{th} \left\|\Xi\right\|_0$, resulting in:

$$\Xi^* = \arg\min_{\Xi} \left(\left\|X' - \Theta(X,U)\Xi\right\|_2 + \lambda \left\|\Xi\right\|_2 + \xi_{th} \left\|\Xi\right\|_0 \right), \tag{3.6}$$

where, Ξ^* are the optimal coefficients and λ and ξ_{th} are the regularization and threshold hyperparameters, respectively. In (3.6), the 0-norm promotes sparsity by eliminating components with low magnitudes ($\xi_i < \xi_{th} \rightarrow \xi_i = 0$), while the $L2$-norm associated with λ regulates the coefficients and promotes small values. The optimal trade-off between model accuracy and sparsity is achieved by fine-tuning the associated hyperparameters. The sparse regression algorithm iteratively applies ridge regression followed by thresholding, effectively removing insignificant terms from Θ, as shown in Fig. 3.1. In each iteration, terms with nonzero coefficients are retained, and ridge regression is recalculated. This process continues until the nonzero coefficients converge or the maximum iteration limit is reached, ensuring a sparse and stable solution.

It's important to note that STRidge works better than other regression methods like STLASSO ($\mathcal{P}(\Xi) = \lambda \|\Xi\|_1 + \xi_{th} \|\Xi\|_0$) and sequentially thresholded least-squares (STLS, $\mathcal{P}(\Xi) = \xi_{th} \|\Xi\|_0$) due to correlated features present in the battery's electrochemical processes [10, 14].

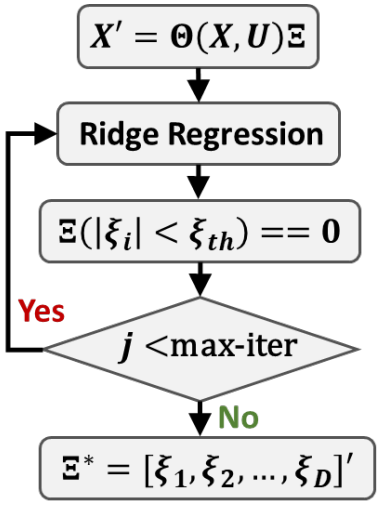

Figure 3.1. Diagram of the STRidge Algorithm.

3.1.2 Hyperparameter Autotuning

The effectiveness of STRidge relies heavily on selecting suitable hyperparameters. These hyperparameters are algorithm-specific properties that govern the learning process and influence the model parameters discovered by the algorithm. With enough data and proper selection of the algorithm's hyperparameters, data-driven learning techniques can achieve impressive performance. However, they often yield suboptimal outcomes due to the vast parameter search space, which contains numerous less-than-ideal solutions. Hyperparameter tuning involves an outer-loop optimization procedure in which an optimization method is applied to fine-tune another optimization [162, 183]. The quality of hyperparameters is

non-deterministic because they cannot be directly derived from mathematical optimization tools; therefore, they must be obtained iteratively. For machine learning applications, automatic hyperparameter tuning algorithms, referred to as autotuners, present an appealing option for automating the training and hyperparameter selection processes.

In this work, a hyperparameter autotuner was created for the STRidge method, as shown in Fig. 3.2. It employs a grid-search strategy to comprehensively explore the search space. Diverging from the conventional approach of using an Akaike Information Criterion (AIC)-inspired cost function to select the threshold parameter [14, 184, 185], which relies solely on training data, this formulation optimizes ξ_{th} by minimizing the cost defined in (3.7).

$$\min_{\Xi} J(\Xi) \stackrel{\text{def}}{=} \rho_1 E_t(\text{SOC}, \widehat{\text{SOC}}) + \rho_2 E_v(\text{SOC}, \widehat{\text{SOC}}) + \rho_3 K, \tag{3.7}$$

This cost incorporates multiple performance metrics, where ρ_1 and ρ_2 penalize prediction errors on the training and validation datasets, respectively, and ρ_3 penalizes model complexity, quantified by the number of terms, K. Previous studies [14, 156, 186] demonstrated that the performance of sparse regression algorithms (e.g., STRidge, STLASSO, STLS) is significantly affected by the threshold parameter ξ_{th}, which plays a crucial role in balancing model accuracy and sparsity. The model's accuracy on training and validation datasets is defined by $E_t(\text{SOC}, \widehat{\text{SOC}})$ and $E_v(\text{SOC}, \widehat{\text{SOC}})$ respectively, and is based on the following relation:

$$E_S(x, \widehat{x}) \stackrel{\text{def}}{=} \text{RMSE}_S(x, \widehat{x}) \stackrel{\text{def}}{=} \sqrt{\frac{\sum_{i=k+1}^{k+m} \left(x[i] - \widehat{x}[i]\right)^2}{m}}, \forall x[i] \in S \tag{3.8}$$

where $\widehat{\mathrm{x}}$ represents the predicted output from the model in the dataset S and m represent the number of elements in S.

The development of the autotuner incorporated the following principles. The STRidge algorithm identifies a nonlinear model (3.4) from a finite collection of samples ($\mathbf{X}$, $\mathbf{U}$) collected from a ground truth distribution $\mathcal{G}_{(\mathbf{x},\mathbf{u})}$ by minimizing an expected loss function $\mathcal{L}$ (3.5). Moreover, the optimality of (3.4) attained through STRidge hinges on the choice of optimal hyperparameters tailored to the selected dataset and library Θ. Optimizing the algorithm's hyperparameters entails minimizing the cost from (3.7). This ensures an optimal tradeoff between model accuracy and complexity. However, in deep learning applications, hyperparameter optimization entails minimizing the expected generalization error (GE), across the search space, $\mathbb{S}$. The theoretical representation of the generalization error ($GE = \mathbb{E}_{(\mathbf{x},\mathbf{u})\sim\mathcal{G}_{(\mathbf{x},\mathbf{u})}}[\mathcal{L}((\mathbf{x},\mathbf{u});\boldsymbol{\Theta}(\mathbf{X},\mathbf{U})\boldsymbol{\Xi}_{\xi_{\mathbf{th}}})]$) quantifies the expected error when applying the model across the full range of possible data values (X, U) [187, 188]. In practice, since the distribution is unknown, the generalization error ($GE = |E_t(SOC, \widehat{SOC}) - E_{cv}(SOC, \widehat{SOC})|$) for a data-driven model is defined as the difference between the empirical loss on the training set and the expected loss on the test (or cross-validation (CV)) set, typically assessed by comparing the errors the model generates on each dataset [183, 189]. This metric provides a measure of the model's ability to generalize effectively from the training data to previously unseen data. It's important to note that since insights gained from unseen data cannot be used to further refine the model, here this evaluation ($E_{cv}(SOC, \widehat{SOC})$) serves as a final validation of the model's generalization capability.

The search algorithm starts with an automated analysis of the magnitudes of a baseline set of non-thresholded coefficients (Ξ) in (3.4), calculated using the Moore-Penrose pseudoinverse. This analysis establishes the upper and lower bounds

for the threshold hyperparameter ξ_{th} search space $\mathbb{S}$, which is represented as a discrete, linearly spaced grid of points between these bounds. The effectiveness of the hyperparameter trial points is evaluated using the model performance metric given in (3.7). Ultimately, the autotuner constructs a model (refer to (3.6)) for each trial point on the grid, identifying ξ_{th}^* associated with the trial point that yielded the optimal model performance.

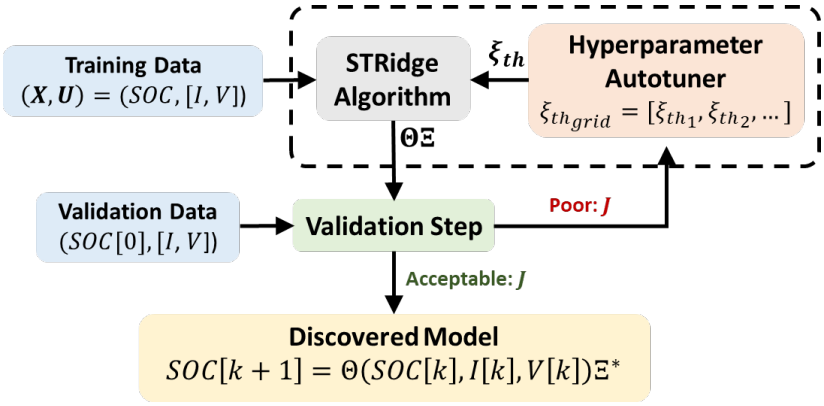

Figure 3.2. Diagram of the Hyperparameter Autotuner.

3.1.3 Monte Carlo Library Search (MCLS)

The sparse nonlinear modeling method operates under the assumption that physical systems are inherently simple, necessitating only a few pertinent terms to represent their dynamics. However, this assumption of sparsity remains valid solely if the function space (feature library) is extensive enough to encompass the distinct dynamic behaviors exhibited by the relevant physical system.

Typically, simple terms like polynomials are employed to construct the feature library. While versatile, these terms fail to capture the complex dynamics of systems like LiBs, which involve various dependencies, including state of charge and temperature. Without prior knowledge of optimal terms for LiBs,

the learning algorithm would need extensive exploration of the function space, potentially leading to intractable problems or inefficient solutions, such as fitting to incorrect nonlinear functions. The feature library is built with explicit physics-informed terms (baseline library Θ_{bl}) to address this shortcoming, enhancing model interpretability and generalizability. Additionally, an automated random search was implemented to efficiently explore the high-dimensional function space, refer to Fig. 3.3. This method iteratively augments Θ_{bl} with randomly chosen nonlinear terms to better capture the complex electrochemical behaviors of LiBs. The augmented library, designated as $\Theta_a = [\Theta_{bl}, \Theta_e(:, \text{idx}_{\text{rand}})]$, includes terms randomly selected (idx_{rand}) from an extended library Θ_e encompassing high-order polynomials, hyperbolic trigonometric functions, and more (e.g., $\Theta_e(X, U) = \begin{bmatrix} \cdots & X^5 & \cdots & U^5 & \cdots & \sinh(U) & \cdots & \cosh(X) & \cdots \end{bmatrix}$).

A Monte Carlo-based random sampling strategy was adopted to choose the supplementary library terms from Θ_e, as it offers an efficient alternative to the conventional manual or brute-force methods for navigating the extensive search space. This Monte Carlo sampling technique generates an array (idx_{rand}) consisting of j random integer values drawn from a discrete uniform distribution spanning the numbers between 1 and D_e. Here, D_e represents the count of terms in Θe, and j is a parameter defined by the user indicating the number of additional terms to be selected. As outlined in §3.1.2, varying model structures (feature libraries Θ) lead to distinct hyperparameter search spaces. Consequently, the hyperparameter autotuner is employed to fine-tune ξ_{th} and the ensuing model coefficients Ξ for each Θ_a derived through the Monte Carlo library search (MCLS) procedure. The number of iterations l is a user-specified parameter, which, along with the desired count of additional terms j, enables extensive customization of the resultant feature library (Θ_a) and the search procedure. Ultimately, the

effectiveness of the identified library Θ_a is evaluated using the same performance metrics as those employed in §3.1.2. The augmented model is represented by:

$$X' = \Theta_a^*(X, U)\Xi^*. \tag{3.9}$$

It's important to note that while initially, Θ_a contains more terms compared to Θ_{bl}, the inclusion of pertinent additional nonlinear terms enables the enhancement of model sparsity. This is achieved by substituting or eliminating combinations of previously required terms through the optimal coefficients Ξ^*.

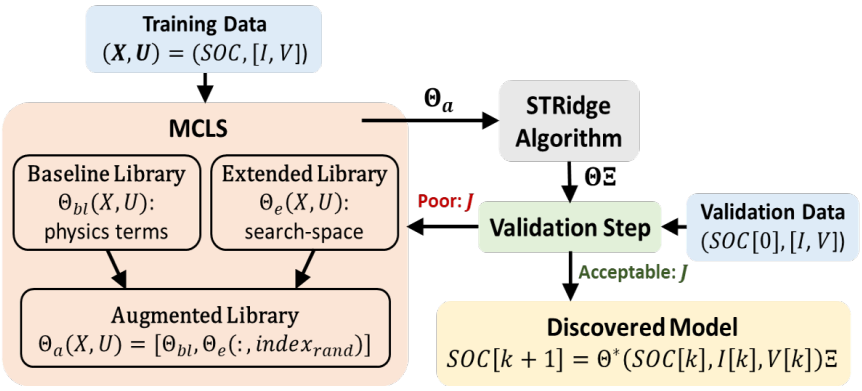

Figure 3.3. Diagram of the Monte Carlo Library Search (MCLS) Algorithm.

3.1.4 Re-calibration of Model Coefficients for Distinct Operating Condition

The extreme temperature conditions that EV energy storage systems encounter necessitate characterizing the battery's temperature-dependent behavior, as fluctuations in temperature affect the battery's available capacity [76]. Therefore, temperature is a crucial and unavoidable factor for SOC methods, and neglecting its influence can result in significant errors [103]. To ensure the efficacy of the forecasting models across the full operational spectrum, encompassing temperatures

from -20°C to 40°C, the methodology is enhanced with a re-calibration approach to optimize model coefficients on new data corresponding to discrete temperature conditions. A constraint-based optimization approach is used to re-calibrate the model coefficients. This process entails minimizing the RMSE-based cost function given in (3.10) using MATLAB's *fmincon* algorithm. (3.10) aims to optimize the model's accuracy in new operating conditions while maintaining the optimal model structure (feature library found via MCLS), refer to Fig. 3.4.

$$\min_{\Xi_{T_i}} . J(\Xi_{T_i}) \stackrel{\text{def}}{=} E_{T_i}(\text{SOC}, \widehat{\text{SOC}}) \tag{3.10}$$

where E_{T_i} is the prediction error on the re-calibration set and T_i is the temperature label (e.g., 25°C) for the re-calibration set. Each new set of coefficients is found through sequential iterations of the optimization routine, initializing the search with a coarse resolution of the coefficient search space with a set maximum allowable deviation from the optimal coefficients of the base model developed at the standard operating condition (e.g., temperature of 25°C). The base model is presented in §5.4. During each iteration, the search space is narrowed as the focus is shifted to areas of high performance while simultaneously increasing the resolution to fine-tune the optimal coefficients that maximize accuracy.

This approach produces a look-up table (LUT) of temperature-dependent model coefficients, extending the models' operational range. It allows for accurate SOC forecasting across varying operating conditions, including the highly nonlinear, low-temperature, low-SOC regime, without increasing model complexity by maintaining the same model structure.

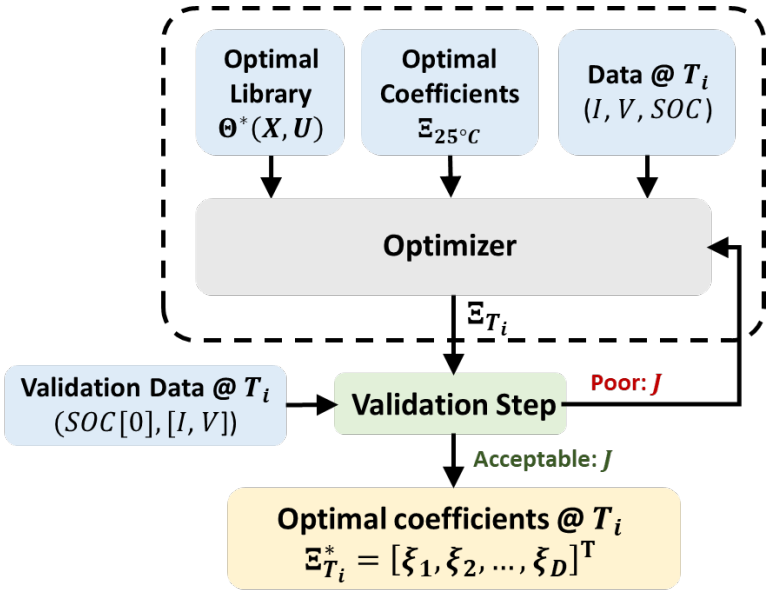

Figure 3.4. Diagram of the Coefficient Re-calibration Algorithm.

3.1.5 Battery Digital Twin of State of Charge Dynamics

The battery digital twin was developed using the methods elaborated in §3.1. It comprises a reduced-order nonlinear data-driven model of the LiB's state of charge (SOC) dynamics. (3.11) illustrates a discrete-time representation of the model, where current I and voltage V act as inputs, and SOC serves as the output.

$$SOC[k+1] = \Theta(SOC[k], I[k], V[k])\Xi \tag{3.11}$$

Physics-informed Library

Battery data can be directly used to predict SOC through the library Θ, enabling real-time forecasting. However, (3.11) depends on properly selecting library terms, especially for complex energy storage systems. Using generic nonlinear terms may result in adequate prediction performance but risks fitting the data into an incorrect nonlinear model, leading to poor performance in unseen

scenarios. To mitigate this, terms related to the fundamental physics governing LiBs are incorporated instead of adding numerous generic terms for general nonlinearities, such as polynomials and mixing terms. The physics-informed terms, including trigonometric, integral, and exponential functions, are derived from the DFN model presented in §4.1.2 to enhance interpretability, generalizability, and computational efficiency. The trigonometric and exponential terms are derived from solid (4.1) and electrolyte (4.2) concentrations, reflecting diffusion processes. These processes involve exponential and zero-order Bessel functions, which can be expressed as trigonometric terms via Fourier transformation. Furthermore, the Butler-Volmer equation (4.3), which delineates the voltage-current relationship at the solid electrolyte interface (SEI) layer, has hyperbolic sine functions in the solution, which are exponential terms. The integral term is inspired by the enhanced coulomb counting method, which is commonly employed to estimate the state of charge with the integral of current in the following form:

$$SOC(t) = SOC(0) - \frac{1}{Q_n} \int_0^t \eta_c I d\tau, \tag{3.12}$$

where Q_n is the battery capacity and η_c is the coulomb efficiency. The Coulomb efficiency η_c requires periodic recalibration to maintain accurate SOC estimations. This is achieved by analyzing the charge extracted from and input into the battery during full charge/discharge cycles. Recalibration is also necessary to account for variations in temperature, battery health, C-rate, and sensor drift.

The enriched baseline library Θ_{bl} included the following terms:

- Terms for general nonlinearities:
  - P: Polynomial (e.g., V^2, ..., I^2, ...)
  - M: Mixing (e.g., $V \cdot SOC$, $V \cdot I$, ...)

- FE: Fractional exponent (e.g., $V^{1.1}$, ..., $I^{2.2}$)

- physics-informed terms

  - T: Trigonometric (e.g., $\sin(V)$, $\cos(V)$, ...)

  - Exp: Exponential (e.g., e^V, e^I, e^{SOC})

  - Int: Integral (e.g., $\int(I)dt$)

The library structure is as follows:

$$\Theta(SOC, I, V) = \begin{bmatrix} | & | & | & | & | & | & | & | & | \\ 1 & I^n & \cdots & I^n V^n & \sin(I) & \cdots & e^I & \cdots & \int I dt \\ | & | & | & | & | & | & | & | & | \end{bmatrix}, \tag{3.13}$$

where each column represents a candidate term and consists of time-series data, such as the current signal with m samples $I \overset{\text{def}}{=} I[k, m]$. The ultimate baseline library was crafted based on the study presented in §5.1, which evaluated the contribution of each term in accurately describing the battery data.

Model Development

The development of the battery digital twin involves training, validation, and cross-validation processes. The training method utilizes the enhanced methodology detailed in section 3.1. This includes employing Monte Carlo library search (MCLS) to determine the most suitable feature library and fine-tuning the hyperparameters and resulting model coefficients through a sparse regression algorithm (STRidge). This process aims to identify coefficients that achieve the optimal balance between accuracy and sparsity. During the training phase, the algorithm has access to the complete input/output time series signals. Then, a validation process is carried out to assess the identified model's predictive accuracy, composed of the

feature library Θ and coefficients Ξ. Here, the model inputs, V and I are used along with the initial SOC condition (SOC[0]), to predict SOC for the entire time series. The model performance, consisting of the accuracy of these predictions and the model complexity (number of terms), is evaluated using the cost function given in (3.7). This stage is pivotal for optimizing the hyperparameters of the STRidge algorithm. This is accomplished by the autotuner (refer to §3.1.2), which modifies the identified hyperparameter values based on minimizing (3.7). Once satisfactory performance is achieved, the model's generalization capability is assessed through a cross-validation process. Similar to the validation phase, the model receives the initial SOC and the inputs, but these inputs come from data not used during training. This demonstrates the model's ability to perform well and adapt to novel conditions.

The modeling approach was applied to both simulated data generated by the DFN model and experimental data outlined in §4.1.2. The model was developed on simulated data, which involved training and validation datasets based on the EPA city driving cycle UDDS and a cross-validation dataset derived from the EPA highway driving cycle US06. The simulated battery datasets were also utilized for the studies on feature library optimization §5.1, sampling rate optimization §5.2, and pulse relaxation §5.3. The ultimate model was constructed utilizing this modeling approach and experimental battery measurements. In this phase, the training and validation datasets are modeled after a stochastic drive cycle developed internally, drawing inspiration from the EPA driving cycles. Meanwhile, the cross-validation data aligns with battery measurements obtained during the US06 drive cycle. Furthermore, the efficacy of the model was expanded to a wide operating range using a stochastic cycle conducted at discrete temperature

conditions ranging from -20°C to 40°C through the development of a tailored look-up table of temperature-dependent model coefficients (refer to § 3.1.4).

3.2 Direct Data-driven Control for Battery Fast-Charging

In this work, a Jacobian learning method is developed to advance fast-charging strategies by significantly reducing battery charging time while minimizing the negative effects on the battery's health. This method optimizes the charging profile by utilizing simulated battery response data generated from the full-order electrochemical model, as detailed in §4.1.2. Importantly, the framework is versatile and can incorporate alternative modeling approaches or direct battery operando measurements. This flexibility allows the solutions to be tailored to the specific mechanical, electrical, and health state of the battery, ensuring greater adaptability and relevance to real-world applications. Moreover, by incorporating the detailed dynamics captured by the model, the approach ensures that the battery is not subjected to unsafe operating conditions, such as exceeding voltage, current, or temperature limits, which could otherwise lead to accelerated degradation or safety hazards.

The Jacobian learning method operates by iteratively refining a charging profile through an adaptive optimization framework. It begins with a sub-optimal baseline solution, such as a standard constant-current constant-voltage charging protocol, to ensure the existence of feasible solutions. At each iteration, the optimizer uses gradient information (Jacobian) from the electrochemical model to guide the adjustment of the charging profile. This adaptive approach progressively improves the initial solution, enhancing performance metrics such as charging speed while preserving long-term battery health.

A key feature of the Jacobian learning method is its ability to account for nonlinearities and complex interactions in battery behavior, which are often overlooked by traditional heuristic charging strategies. By systematically incorporating the battery's dynamic response into the optimization process, the method ensures robustness across a wide range of operating conditions, including varying temperatures and states of charge. Additionally, the use of model-informed learning mitigates the risk of overfitting to specific scenarios, allowing for broader applicability and real-world viability.

This approach represents a significant advancement in the design of fast-charging protocols. Unlike conventional methods, which often rely on fixed charging profiles or empirical adjustments, the Jacobian learning method dynamically tailors the charging process to the battery's evolving state. As a result, it provides an efficient, adaptive, and physically grounded solution that balances the competing objectives of rapid energy delivery and long-term durability.

3.2.1 Problem Formulation

The complex nonlinear lithium-ion battery systems are assumed to have the following general form:

$$\dot{x} = f(x, u). \tag{3.14}$$

where x represents the states and u represents the inputs. The goal is to find an optimal input (charging profile), u^*, that maximizes SOC within a set charging duration. The general formulation of the optimization criteria is given by:

$$u^* = \arg\min_{u \in U} \int_0^{t_f} \varphi(x(t), u(t), t) dt \tag{3.15}$$

subject to the constraints:

$$u_{lb} \leq u(t) \leq u_{ub}$$

$$x_{lb} \leq x(t) \leq x_{ub}$$

where the constraints are defined by lower and upper bounds on the inputs (u_{lb}, u_{ub}) and the states (x_{lb}, x_{ub}), respectively. The constraints include i) bounds on the electrical current input and ii) limitations on the state variables for safe LiB operation, such as a maximum and minimum battery voltage and a maximum temperature. These constraints are summarized in Table 3.2. It's important to note that the specific optimization criteria and constraints depend on the specific battery and the desired charging strategy. Furthermore, this method is amenable to changes, e.g., other criteria and constraints.

	Optim. Criteria / Constraints
SOC [%]	Maximize value at t_f
SOC_d [%]	100% (fully charged)
Temperature [$^\circ C$]	maximum: 57
Voltage [V]	maximum: 4.2
	minimum: 2.5
Current [C-rate]	maximum: 2.5
	minimum: 0

Table 3.2. Optimization Criteria and Constraints

3.2.2 Jacobian Learning Optimization

This section outlines the methodology of Jacobian learning (JL), an adaptive optimization approach. JL employs learning techniques to identify and recursively update the system's input/output sensitivity. It is leveraged to discern the

dominant characteristics of the target system using input-output data, enabling model-free control of complex systems.

The Jacobian learning process is executed through a recursive least squares approach [112]. After this process is completed and the Jacobian (input-output sensitivity) is acquired, it is applied with a gradient-descent optimization strategy to perform constrained optimizations of the inputs. The optimization process contains two sequential steps: the first step entails conducting a full continuous-time simulation, while the second step (discrete time) utilizes insights from the first step to map out the subsequent simulation [179, 190]. This iterative process persists until the optimization metric (e.g., SOC) converges to the optimal solution or until a predefined maximum iteration limit is reached.

JL Problem Formulation Here, a class of static models is introduced, commonly used for processes that evolve slowly or systems where dynamic effects can be neglected relative to the sampling rate. These models are assumed to be zero-order and are nonlinear but smooth [112]. As depicted in Fig. 3.5, the desired output y_d is attained through an iterative optimization of the inputs u to the plant, based on a data-driven model (Jacobian) developed from measurements of the inputs u and outputs y. An overview of the methodology is given below.

The inputs $u[k]$ and outputs $y[k]$ are assumed to be related with a static nonlinearity

$$y[k] = S(u[k]) \tag{3.16}$$

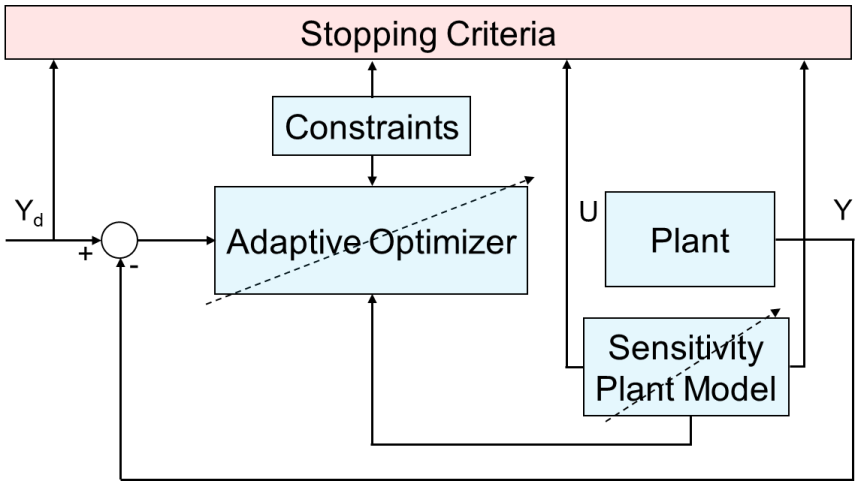

Figure 3.5. Diagram of the Learning and Optimization Algorithm

where $S(u[k])$ is a nonlinear and smooth function. Here, the controller aims to minimize the error, $e[k]$, between the system output, $y[k]$, and the desired output,y_d, by optimizing the input vector $u[k]$. The error is given by:

$$\|e[k]\|_2 = \|y[k] - y_d[k]\|_2, \tag{3.17}$$

where $y_d[k]$ is the desired output vector at time k.

Next, the recursive least squares approach for learning the Jacobian $\mathbb{J}[k]$ is discussed. This approach is also used to recursively update $\mathbb{J}[k]$ to maintain the learned sensitivity. First, a linearized time-varying approximation of the mapping S is considered:

$$\Delta y[k] = \mathbb{J}[k]\Delta u[k] \tag{3.18}$$

where,

$$\begin{aligned} \Delta u[k] &= u[k] - u[k-1] \quad , & \Delta u[k] &\in R^r \\ \Delta y[k] &= y[k] - y[k-1] \quad , & \Delta y[k] &\in R^q \end{aligned} \tag{3.19}$$

and r and q represent the number of inputs and outputs, respectively. It's important to note that for multi-output systems ($q > 1$), (3.18) is decompose into q single-output subsystems $\mathbb{J}_j[k]$:

$$\Delta y_j[k] = \mathbb{J}_j[k] \Delta u[k] \tag{3.20}$$

where $j = 1, 2, ..., q$. The optimal input update for minimizing the cost in (3.17) when the Jacobian $\mathbb{J}[k]$ is known can be found from (3.18) using the pseudo-inverse $\mathbb{J}^\dagger[k]$ as [179].

$$u[k+1] = u[k] + \mathbb{J}^\dagger[k](y_d[k] - y[k]) \tag{3.21}$$

To prevent singularities, regularization of $\mathbb{J}^\dagger[k]$ is performed using $H[k]$ as:

$$\begin{aligned} H[k] &= \mathbb{J}^T[k](\mathbb{J}[k]\mathbb{J}^T[k] + \rho I_q)^{-1} \quad \text{for} \quad r \geq q \\ H[k] &= (\mathbb{J}^T[k]\mathbb{J}[k] + \rho I_q)^{-1}\mathbb{J}^T[k] \quad \text{for} \quad r \leq q \end{aligned} \tag{3.22}$$

where I_q is the $q \times q$ identity matrix and ρ is a small positive constant ($\rho \in (0, 1)$). However, often, the Jacobian is unknown and must be estimated. Once learned, the estimated Jacobian $\widehat{\mathbb{J}}$ can be used as a feedback control law of the form:

$$u[k+1] = u[k] + \widehat{H}[k] G(y_d - y[k]) \tag{3.23}$$

where G represents the control gains with its diagonal elements as $g_i \in (0, 2)$. If $r \leq q$, $\widehat{H}[k]$ is defined as:

$$\widehat{H}[k] = (\widehat{\mathbb{J}}^T[k]\widehat{\mathbb{J}}[k] + \rho I_q)^{-1}\ G\ \widehat{\mathbb{J}}^T[k] \tag{3.24}$$

Here, an adaptive learning approach is employed to find and recursively update the Jacobian to account for Jacobian changes in time, which can be represented by:

$$\mathbb{J}_j[k+1] = \mathbb{J}[k] + w_j[k], \tag{3.25}$$

$$\Delta y_j[k] = \Delta u^T[k]\mathbb{J}_j[k] + v_j[k],\ j = [1, q], \tag{3.26}$$

where the vector $w_j[k]$ signifies the process noise while $Q_j = E\{w_j^T[k]w_j[k]\}$ denotes the expected covariance of the model's imprecision, and $v_j[k]$ characterizes the measurement noise while $R_j = E\{v_j^2[k]\}$ represents the expected variance of the measurement noise [190]. Further, the Jacobian of each simplified subsystems (3.20) can be estimated as:

$$\widehat{\mathbb{J}}_j^T[k] = \widehat{\mathbb{J}}_j^T[k-1] + \frac{P_j[k-1]\Delta u[k](\Delta y[k] - \widehat{\mathbb{J}}_j[k-1]\Delta u[k])}{R_j + \Delta u^T[k]P_j[k-1]\Delta u[k]} \tag{3.27}$$

$$P_j[k] = P_j[k-1] - \frac{P_j[k-1]\Delta \mathrm{u}[k]\Delta u^T[k]P_j[k-1]}{R_j + \Delta u^T[k]P_j[k-1]\Delta u[k]} + Q_j. \tag{3.28}$$

To accommodate for constraints, the optimal feedback control law (3.23) can be reformulated as a constraint-based optimization problem given by:

$$u^*[k] = \underset{u[k]}{\arg\min}(\|y_d - \widehat{y}[k]\|^2 + \gamma\|u[k] - u[k-1]\|^2) \tag{3.29}$$

s.t. $\widehat{y}[k] = y[k-1] + \widehat{\mathbb{J}}[k](u[k] - u[k-1])$.

The update in (3.29) can be implemented through widely accessible quadratic programming solvers, such as Matlab's FMINCON and LSQLIN functions.

4. BATTERY DATA GENERATION AND COLLECTION: SIMULATED AND EXPERIMENTAL METHODS

4.1 Battery Data for Digital Twining Process

Using simulated and experimental data, we developed and tested our data-driven methodology and the physics-informed battery digital twins. The processes for generating and collecting simulated battery data for the modeling tasks is detailed in §4.1.1, while the experimental procedures are covered in §4.1.3.

4.1.1 Battery Simulations

Simulations were conducted using the DFN model detailed in §4.1.2, implemented in the Python Battery Mathematical Modeling (PyBaMM) framework [33], where we tested a commercial LGM50 21700 cylindrical cell with a capacity of 5 Ah. This cell consists of Nickel Manganese Cobalt Oxide (NMC) 811 as the positive electrode and bi-component graphite (SiO_x) as the negative electrode [25], with charge and discharge cutoff voltages set at 4.2 V and 2.5 V, respectively. The maximum continuous charging current is 1.44 A (0.3 C-rates) with a charge cutoff current of 50 mA. For our training/validation and cross-validation, we utilized the current signal (battery input) and corresponding voltage and SOC signals (battery outputs) from standard EPA (U.S. Environmental Protection Agency) drive cycles, specifically the UDDS (Urban Dynamometer Driving Schedule) city driving cycle and the US06 highway driving cycle. We note that mapping the drive cycles to electrical current profiles entails converting the vehicle's speed

and acceleration data from the drive cycle into the power demands of the electric drivetrain. These power demands are subsequently used to determine the corresponding current drawn from the battery. In this work, we utilize the mappings developed in [27, 33].

The simulations were initialized with a SOC of 90% and conducted at an ambient/cell temperature of 25°C. The schematic in Fig. 4.1 illustrates the process of simulating the battery's response and collecting the input/output data.

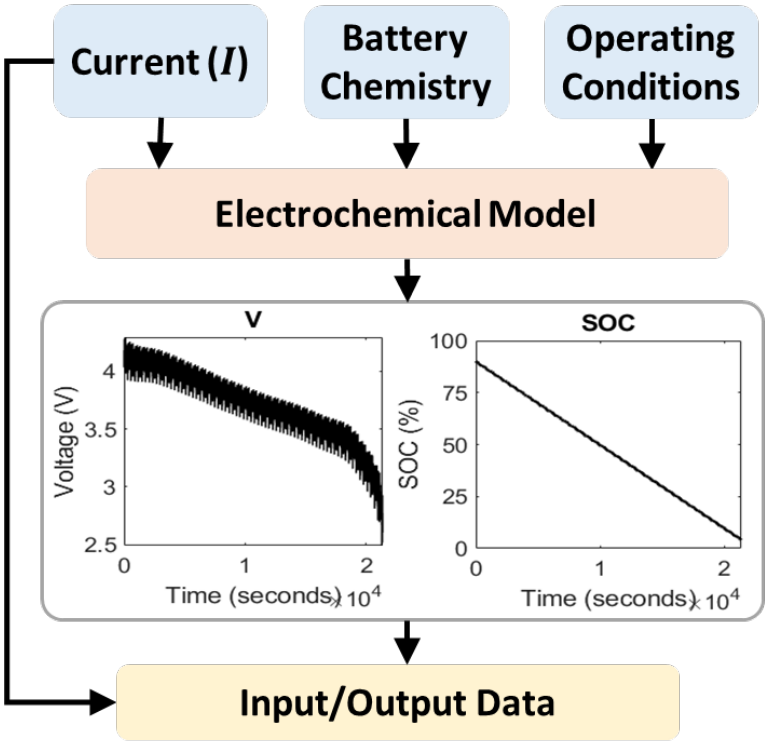

Figure 4.1. Diagram of Data Collection Process.

4.1.2 Physics-based Model

The simulated data originated from the application of the Doyle Fuller-Newman (DFN) physical battery model [18]. The Doyle-Fuller-Newman model (a.k.a Pseudo-two-Dimensional (P2D) model) presents a comprehensive electro-

chemical Li-ion battery model that describes the battery's internal processes, including diffusion, intercalation and electrochemical kinetics, using principles derived from porous electrode and concentrated solution theories [35, 61].

DFN comprises a separator, two electrically separated porous electrodes, anode and cathode, respectively, and the electrolyte as shown in Figure 2.1. Lithium exists in two phases: solid in the electrode material (anode and cathode) and liquid when dissolved in the electrolyte. Li-ions in the solid phase are transported by a diffusion process within the active material along the r-axis. This process, which is related to the lithium concentration in the solid phase $c_s^{\pm}(x, r, t)$, can be modeled using radially symmetric diffusion in spherical coordinates as in:

$$\frac{\partial c_s^{\pm}}{\partial t}(x, r, t) = \frac{1}{r^2}\frac{\partial}{\partial r}\left[D_s^{\pm} r^2 \frac{\partial c_s^{\pm}}{\partial r}(x, r, t)\right] \tag{4.1}$$

where the superscript "$\pm$" indicates the positive and negative electrodes, $D_s^{\pm}$ represents the diffusion coefficient, and r and x denote the radial and longitudinal directions, respectively. In the liquid phase, the Li-ions migrate along the x-axis to the opposite electrode through the solid-electrolyte interphase via Butler-Volmer kinetics. Here, the electrolyte concentration $c_e(x, t)$ is modeled using a combination of Fick's law of linear diffusion and molar flux $j_n^{\pm}(x, t)$ as follows:

$$\frac{\partial c_e}{\partial t}(x, t) = \frac{\partial}{\partial x}\left[D_e \frac{\partial c_e}{\partial x}(x, t)\right] + \frac{(1 - t_c^0)a^{\pm}}{\varepsilon_e F} j_n^{\pm}(x, t), \tag{4.2}$$

where ε_e is the volume fraction of the electrolyte, F is the Faraday's constant, t_c^0 is the transference number and $a^{\pm}$ is the specific interfacial surface area. D_e, the

diffusion coefficient, is a function of electrolyte concentration. The molar flux is given by the Butler-Volmer equation:

$$j_n^{\pm}(x,t) = \frac{1}{F} i_0^{\pm}(x,t) \left[e^{\frac{\alpha_a F}{RT}\eta^{\pm}(x,t)} - e^{-\frac{\alpha_c F}{RT}\eta^{\pm}(x,t)} \right], \tag{4.3}$$

where α_a and α_c are anodic and cathodic charge transfer coefficients, respectively, R is the universal gas constant, and T is the temperature. The exchange current density $i_0^{\pm}(x,t)$ and the intercalation over-potential $\eta^{\pm}(x,t)$ are given by:

$$i_0^{\pm}(x,t) = k^{\pm} \left[c_{ss}^{\pm}(x,t) \right]^{\alpha_c} \left[c_e(x,t) \left(c_{s,\max}^{\pm} - c_{ss}^{\pm}(x,t) \right) \right]^{\alpha_a} \tag{4.4}$$

$$\eta^{\pm}(x,t) = \phi_s^{\pm}(x,t) - \phi_e(x,t) - U^{\pm}(c_{ss}^{\pm}(x,t)) - F R_f^{\pm} j_n^{\pm}(x,t) \tag{4.5}$$

where the solid phase surface concentration is defined as $c_{ss}^{\pm}(x,t) = c_s^{\pm}(x, R_s^{\pm}, t)$ while $c_{s,\max}$ is the maximum possible concentration in the solid phase. $U^{\pm}$ represents the open-circuit potential and $R_f^{\pm}$ the solid-electrolyte interphase film resistance. $\phi_s^{\pm}(x,t)$ and $\phi_e(x,t)$ represent the electric potential in the solid (4.6) and electrolyte phases (4.7), respectively.

$$\frac{\partial \phi_s^{\pm}}{\partial x}(x,t) = \frac{i_e^{\pm}(x,t) - I(t)}{\sigma^{\pm}} \tag{4.6}$$

$$\frac{\partial \phi_e}{\partial x}(x,t) = \frac{-i_e^{\pm}(x,t)}{\kappa} + \frac{2RT}{F}(1 - t_c^0) \times \left(1 + \frac{d \ \ln \ f_{c/a}}{d \ \ln \ c_e}(x,t) \right) \frac{\partial \ \ln \ c_e}{\partial x}(x,t) \tag{4.7}$$

Here, $i_e^{\pm}$ is the ionic current, $I(t)$ is the applied current density, $f_{c/a}$ is the mean molar activity coefficient in the electrolyte, and $\sigma^{\pm}$ and κ are the solid and electrolyte conductivity, respectively. Moreover, κ and $f_{c/a}$ are functions

of electrolyte concentration $c_e(x,t)$. Voltage $V(t)$ is the difference in the solid potential $\phi_s^{\pm}$ between the two ends of the electrode, as follows:

$$V(t) = \phi_s^+(0^+,t) - \phi_s^-(0^-,t). \tag{4.8}$$

The battery's available energy can be determined by the volume-averaged solid-phase lithium concentration in the anode [175]. This calculation assumes the anode capacity to be the limiting factor and yields the calculation of the state of charge (SOC) as follows:

$$\mathrm{SOC}(t) = 100 \left[\frac{\left(\frac{1}{L^- \bar{c}_{s,max}^-} \int_0^{L^-} c_{s,avg}^-(x,t)dx \right) - \theta_{min}}{\theta_{max} - \theta_{min}} \right] \tag{4.9}$$

where $c_{s,avg}^-$ represents the volume-averaged solid phase concentration in each solid particle in the anode, $c_{s,max}^-$, represents the maximum solid phase concentration in the anode, L^- represents the anode length, and θ_{max} and θ_{min} represent the SOC at the fully charged/discharged states, respectively. We note that these parameters are defined by the anode's stoichiometric limits. Additionally, the standard DFN model is extended with a thermal model that couples the porous electrode theory with an energy conservation approach to describe the cell's thermal behavior, including Ohmic heating in both the solid and the electrolyte, as well as reversible and irreversible heating resulting from electrochemical reactions [28]. The spatially averaged cell temperature ($\bar{T}$) is given by:

$$\bar{T}(t) = \frac{1}{L} \int_0^L T(x,t)dx. \tag{4.10}$$

This complex battery model provides accurate information about the internal states, such as SOC, across various operating conditions. However, its formulation

leads to a large, computationally intensive model. The coupled partial differential equations (PDEs) defining the DFN model are too resource-demanding to be executed in real-time by the hardware typically available in a BMS [8]. Furthermore, the requirement for detailed information on the battery's composition and internal parameters suggests that it is best suited for setting performance and efficiency benchmarks or offline applications. In this work, we employed the DFN model for offline data generation and benchmarking our physics-informed battery digital twin, reducing complexity and relaxing the need for detailed cell composition knowledge while maintaining production accuracy.

4.1.3 Experiments

Battery experiments were conducted on a cell mirroring the model, using commercially available battery testers and a thermal test chamber to regulate temperature. The experimental setup is illustrated in Fig. 4.2. Initially, the cell was subjected to multiple charge and discharge cycles using the constant current constant voltage (CCCV) method. Then, the battery was fully charged using CCCV at a 0.3 C-rate per the manufacturer's recommendations, followed by a two-hour rest period to allow the cell to reach a steady state. Subsequently, a stochastic current profile was applied until the lower voltage limit was reached. Similar steps were repeated for experiments corresponding to the US06 and UDDS cycles. The EPA and stochastic current profiles are shown in Fig. 4.3. Since the SOC is not directly measured in the experiments, we utilized an enhanced Coulomb counting approach [191, 192] based on the battery's Coulombic efficiency (η_c). η_c represents the ratio of charge withdrawn from the battery to the charge injected into it over a full cycle. We recalibrated η_c for each experiment to ensure accurate SOC references. However, we acknowledge that such frequent recalibration is

Table 4.1. Electrochemical Modeling Nomenclature

Symbols	**Description [units]**
$c_s^{\pm}$	Li concentration in solid phase [mol/m^3]
c_e	Li concentration in electrolyte phase [mol/m^3]
$\phi_s^{\pm}$, ϕ_e	Solid, electrolyte electric potential [V]
$i_e^{\pm}$	Ionic current [A/m^2]
$j_n^{\pm}$	Molar ion flux [mol/m^2-s]
$i_0^{\pm}$	Exchange current density [A/m^2]
$\eta^{\pm}$	Overpotential [V]
$c_{ss}^{\pm}$	Li concentration at solid particle surface [mol/m^3]
I	Applied current [A/m^2]
V	Terminal voltage [V]
$D_s^{\pm}$, D_e	Diffusivity of solid, electrolyte phase [m^2/s]
t_c^0	Transference number [-]
$\varepsilon_s^{\pm}$, ε_e	Volume fraction of solid, electrolyte phase [-]
F	Faraday's constant [C/mol]
$\sigma^{\pm}$	Conductivity of solid [1/Ω-m]
κ	Conductivity of electrolyte [1/Ω-m]
R	Universal gas constant [J/mol-K]
T	Temperature [K]
$f_{c/a}$	Mean molar activity coefficient in electrolyte [-]
$a^{\pm}$	Specific interfacial surface area [m^2/m^3]
α_a, α_c	Anodic, cathodic charge transfer coefficient [-]
$k^{\pm}$	Kinetic reaction rate [(A/m^2)(mol^3/mol)$^{(1+\alpha)}$]
$c_{s,\max}^{\pm}$	Maximum concentration of solid material [mol/m^3]
$U^{\pm}$	Open circuit potential of solid material [V]
$R_f^{\pm}$	Solid-electrolyte interphase film resistance [Ω-m^2]
$R_s^{\pm}$	Particle radius in solid phase [m]

impractical and primarily limited to laboratory settings. A thermal test chamber was employed to ensure stable and consistent temperatures during the battery experiments. The chamber regulated the ambient temperature while monitoring both ambient and cell surface temperatures using dedicated probes. Experiments at varying temperatures began only after the cell surface temperature equilibrated with the ambient temperature. The testing was conducted at discrete temperature points within the range of -20°C to 40°C. Fig. 4.4 illustrates the SOC references corresponding to different temperature conditions used in this work.

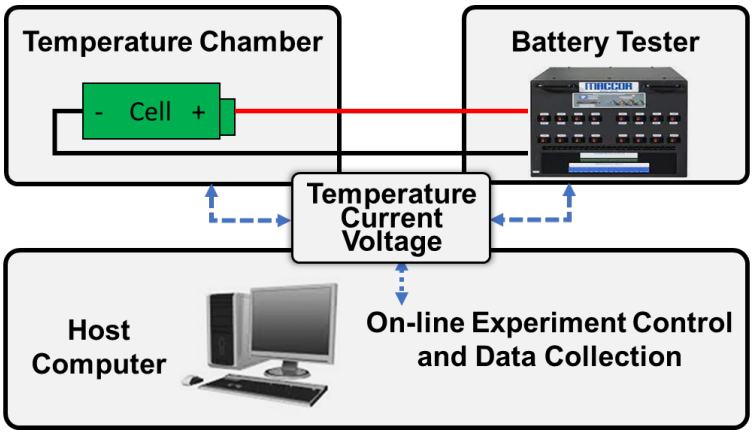

Figure 4.2. Diagram of Experimental Setup.

4.2 Battery Data for Fast Charging

Similar to §4.1, we employed the Python Battery Mathematical Modeling (PyBaMM) framework [33] for efficient battery simulations. Additionally, we utilized PyBaMM's "plug-and-play physics" methodology, to facilitate the integration of thermal effects into the DFN model [28].

In this study, we adopt the battery chemistry NMC 811 LGM50 21700 cylindrical cell. [25]. We used the PyBaMM-DFN model for the calculation of relevant battery outputs, such as voltage V, state-of-charge SOC, and temperature T,

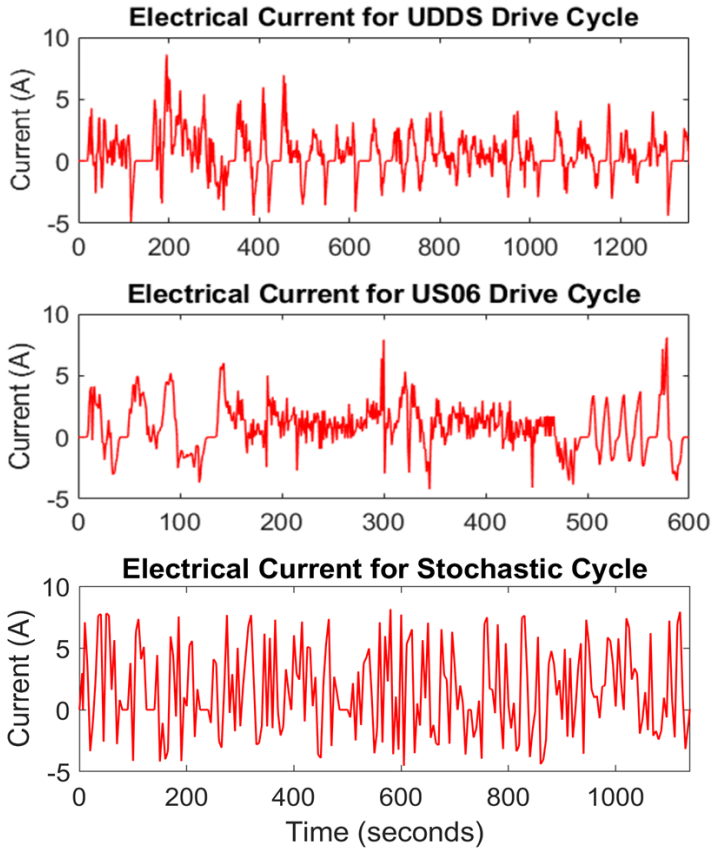

Figure 4.3. Current (I) for UDDS, US06 and stochastic driving cycles.

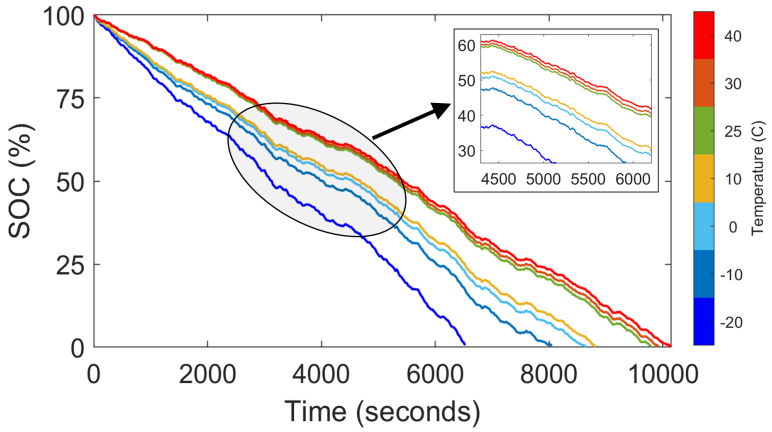

Figure 4.4. SOC References for Discrete Temperature Conditions.

for a selected electrical current I input (e.g., CC-CV, pulse charging techniques, etc.). The battery simulation and data collection process is shown in Fig. 4.5.

We note that the battery simulations initialize the battery SOC to 0% and use an ambient/initial-cell temperature of 25^{o}C.

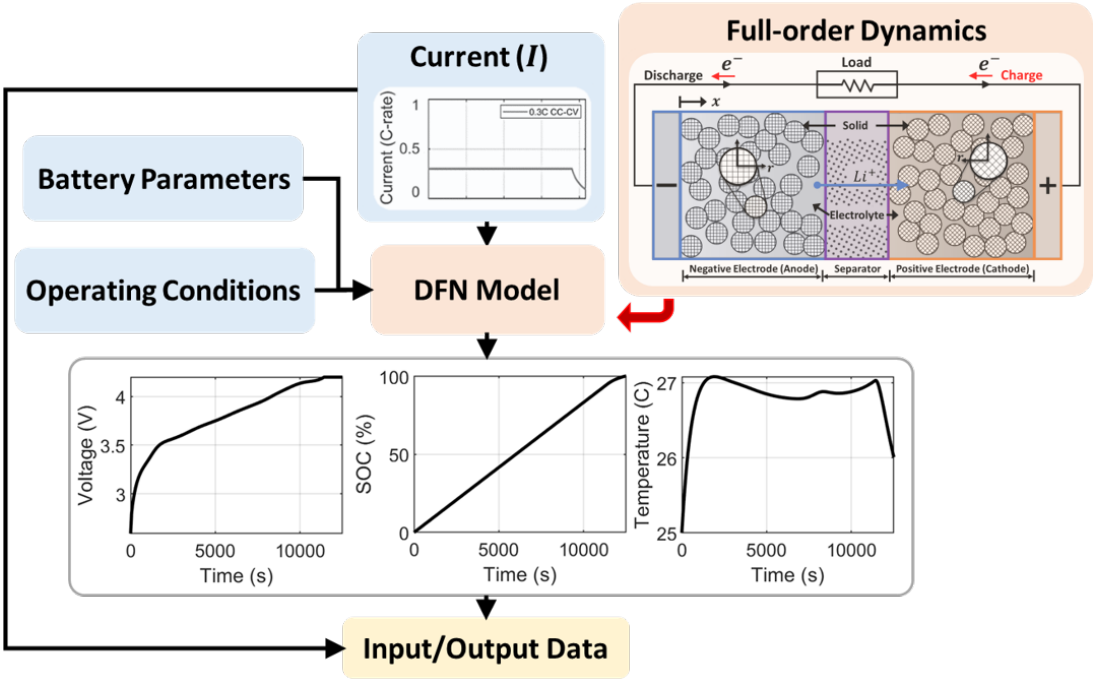

Figure 4.5. Schematic of Data Collection Process.

5. BATTERY DIGITAL TWIN RESULTS

Here, we highlight the effectiveness of the battery digital twin of SOC dynamics introduced in §3.1. Our investigations explored various factors, such as the proper selection of library terms and data sampling rates, all of which have significant implications for the accuracy and sparsity of the resulting model. We assessed the outcomes of these investigations by examining validation RMSE values, which are detailed in §5.1 and §5.2 respectively. The summarized results and visualizations depicting the models' training, validation, and cross-validation processes, developed using both simulated and experimental data, are presented in §5.4.

5.1 Feature Library Optimization

This section presents a study investigating the effects of distinct feature libraries on model performance. Our study introduces domain knowledge to the learning process via our physic-inspired terms and optimizes a baseline library (Θ_{bl}) based on maximizing model sparsity and prediction accuracy. The library comprises linear and nonlinear transformations of the model input (V and I) and output (SOC) variables. Moreover, the resulting Θ_{bl} serves as a stable initialization point for our automated random search of additional library terms (MCLS, see §3.1.3) for further enhancement of the model's performance. The relevance of the library terms used in this study, including our physics-informed terms, is detailed in §3.1.5.

Our study followed these procedures: (i) Simulate the battery response, including V and SOC, for the UDDS current input I conducted at 25°C using the electrochemical model detailed in §4.1.2. (ii) Craft distinct feature libraries, including various combinations of terms such as P, T, etc., using the simulated input/output data. (iii) Train and validate a model for each distinct library. (iv) Evaluated each model's performance based on sparsity and prediction accuracy (3.7). (v) Select the best-performing library as Θ_{bl}.

Our results, briefly summarized in Table 5.1, showed that feature libraries that excluded fractional exponent (FE) terms experienced lower complexity and comparable or higher accuracy than libraries with FE terms. Furthermore, we uncovered that individual inclusion of our physics-informed terms, such as exponential (Exp) and integral (Int) terms, did not improve accuracy, but their joint inclusion led to simpler models and lower prediction errors (RMSE). These findings were used to select relevant terms for the baseline feature library Θ_{bl}, including P, M, T, Exp, and Int terms. The corresponding model $\text{SOC}[k+1] = \Theta_{bl}(\text{SOC}[k], I[k], V[k])\Xi$ and the sparse coefficients Ξ were identified via STRidge on the simulated UDDS data. An RMSE performance of $2.2 \times 10^{-}2$ was achieved with Θ_{bl} consisting of only 26 terms. We note that the selected baseline library was used in subsequent studies, where improved accuracy and sparsity were achieved.

5.2 Sampling Rate Optimization

In this study, we examine the impact of data sampling rate on the performance of our data-driven modeling approach and optimize it to maximize prediction accuracy. We evaluated a range of data sampling rates, from 50 ms to 1000 ms, based on the minimum step size of commercially available battery testers and the

Table 5.1. Feature Library Optimization Results

Library Terms	**RMSE**	**Sparsity**
P, M, T, FE, Exp	6.8	40
P, M, T, FE, Int	6.6	37
P, M, T, FE, Exp, Int	3.5	41
P, M, T, Exp, Int	$2.2 \times 10^{-}2$	26

Polynomial (P), Fractional Exponent (FE), Mixing (M), Trigonometric (T), Exponential (Exp), Integral (Int)

1-second sampling rate of the current inputs for the standard EPA drive cycles. This study comprises two parts. First, we examined the effect of varying the sampling rate of the UDDS discharge cycle while maintaining consistent initial and final SOC conditions for each variant. This resulted in different numbers of samples for each variant: the fastest rate of 50 ms yielded 36,000 samples, whereas the slowest rate of 1 s yielded 1,800 samples. In the second part of the study, we examined the effect of maintaining a consistent sample size (number of data samples) while varying the sampling rate of our UDDS discharge cycle. This was done to determine if changes in model performance were due to variations in sample size, the ability to capture more detailed information about the battery's dynamics with faster sampling rates, or a combination of both factors.

For part one, we followed the procedure below to evaluate the impact of different sampling rates on performance: (i) Resample the UDDS current input (I) for each sampling rate using spline interpolation. (ii) Simulate the battery response (V, SOC) at 25°C, for each resampled I using the DFN model from §4.1.2. (iii) Train and validate a model for each resampled input/output data set using the baseline library Θ_{bl} from §5.1. (iv) Evaluate each model's performance based on prediction errors (RMSE). (v) Choose the best-performing data sampling

rate based on the lowest RMSE. Part two of the study followed similar steps, but instead of the resampling process in step (i), we conducted an under-sampling process to vary the sampling rate while maintaining a consistent sample size. This involved a current input comprising sequential charging and discharging cycles based on the 50 ms UDDS current input. This current input fully discharges and subsequently charges the 5Ah cell in 1.325×10^6 samples. To ensure that the resampled sets contain the same number of samples, we repeated the current input to generate larger datasets that could be under-sampled to achieve the desired sampling rate while still containing information about the full SOC range (0% to 100%). This resulted in 20 resampled datasets covering sampling rates ranging from 50 ms to 1000 ms in 50 ms intervals.

The results for part one are summarized in Table 5.2, while Fig.5.1 presents the results for part two. Our results suggest that models trained on the UDDS data resampled at 50 ms achieved the best performance. This trend was consistent across both studies, with slower sampling rates resulting in worsened RMSE performance. It's important to highlight that all models employed the same feature library (Θ_{bl}), comprising 26 terms, indicating that the enhanced RMSE was solely due to the alteration in sampling rate. Moreover, we infer that the improvement is attributable to both the augmented number of data samples and the faster sampling rate, which provides more intricate details about the battery's dynamics. The latter aspect is further explored in §5.3. Additionally, the formulation of RMSE, which relies on the number of samples utilized, also played a role in enhancing performance.

Table 5.2. Sampling Rate Optimization Results (Part-1, varied sample size)

Sampling Rate	Sample Size	RMSE	Sparsity
50 ms	36000 samples	$5.44 \times 10^{-}6$	26
250 ms	7200 samples	$6.55 \times 10^{-}4$	26
450 ms	4000 samples	$3.52 \times 10^{-}3$	26
650 ms	2770 samples	$8.89 \times 10^{-}3$	26
850 ms	2120 samples	$1.54 \times 10^{-}2$	26
1000 ms	1800 samples	$2.20 \times 10^{-}2$	26

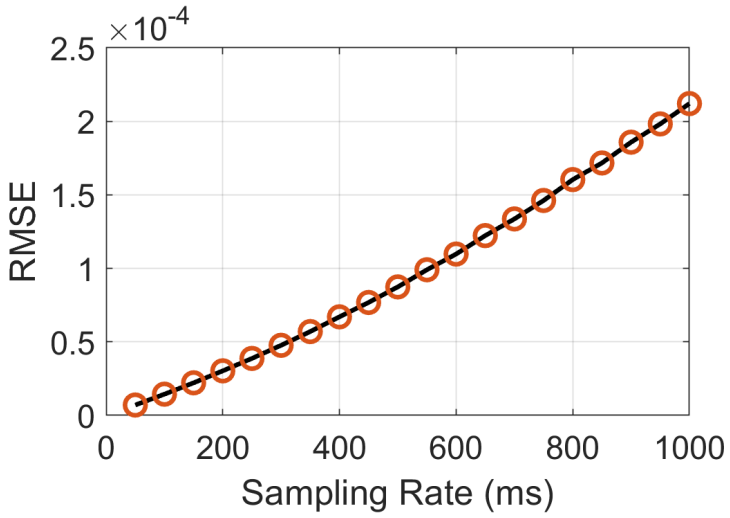

Figure 5.1. Sampling Rate Optimization Results (Part-2, consistent sample size).

5.3 Pulse-Relaxation Study

Here, we delved deeper into determining the optimal sampling rate for Li-ion battery SOC dynamics by examining the battery's response to pulse excitation input. Our investigation involved applying a current input consisting of charge

and discharge pulses and relaxation periods to the electrochemical model detailed in §4.1.2. The charge and discharge pulses utilized a maximum current of 0.05A with a pulse width of 1 second. Following each pulse a rest period was applied, where a current of 0A was applied for 14 seconds to allow the battery to reach a steady state. We analyzed the battery's response throughout the rest periods to uncover the time scale of the SOC dynamics. Our findings, depicted in Fig. 5.2, reveal that these dynamics evolve in the order of milliseconds. Moreover, these results align with the time scale for the interfacial charge transfer kinetics observed in [193]. The interfacial charge transfer is typically assumed to follow Butler–Volmer kinetics and exhibits high SOC dependencies [194]. Therefore, to accurately capture the SOC dynamics from measurement data, it's imperative to collect data using a sampling rate in the order of milliseconds or faster. This confirms the validity of our selection of 50 ms from §5.2 as the optimal sampling rate for LiB applications.

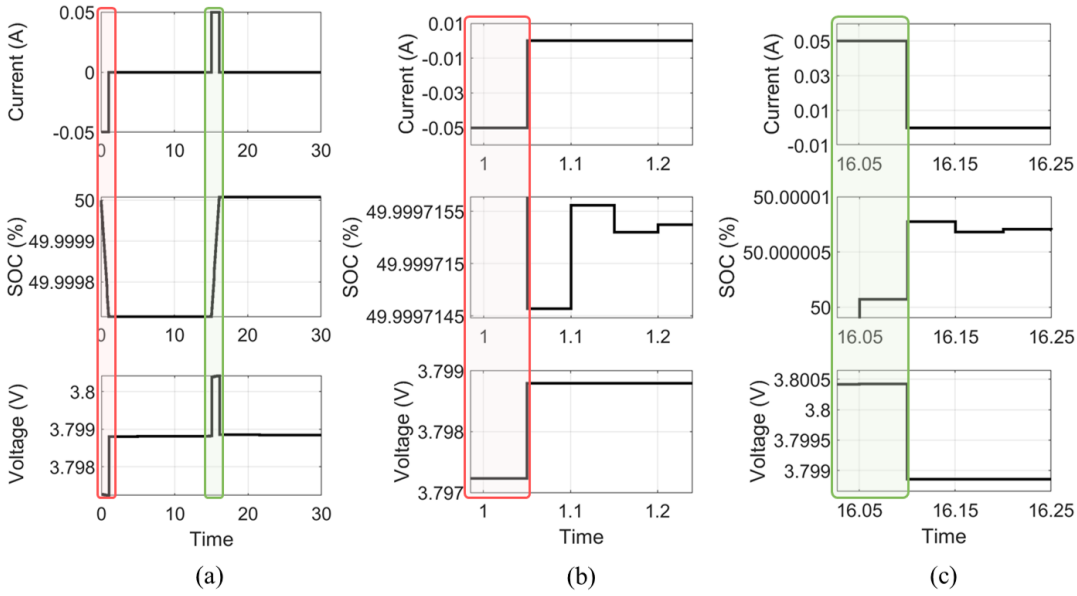

Figure 5.2. Pulse Relaxation Study: (a) Current Input, (b) Discharge Pulse (red) & Relaxation, (c) Charge Pulse (green) & Relaxation.

Table 5.3. Battery Digital Twin of SOC Dynamics Results

Simulated Data	**RMSE**	**Sparsity**
Training	5×10^{-7}	⋮
Validation	1.3×10^{-3}	10
Cross-validation (US06)	1.4×10^{-3}	⋮
Experimental Data	**RMSE**	**Sparsity**
Training	2.2×10^{-6}	⋮
Validation	4.8×10^{-4}	8
Cross-validation (US06)	8.5×10^{-4}	⋮
Library Terms (Θ^*)	SOC, SOC·V, sin(SOC), sin(V), exp(SOC), $\int$(I), sinh(SOC), cosh(V)	

5.4 Physics-informed and Temperature-Dependent Digital Twin of Battery SOC Dynamics

Here, we present the battery digital twin developed with the library terms detailed in §5.1, the re-sampled input/output data from §5.2, and our enhanced data-driven modeling methodology, including auto-tunning of the hyperparameters detailed in §3.1 and our Monte Carlo search of additional nonlinear terms from §3.1.3.

We devised an efficient and precise battery digital twin by constructing a reduced-order nonlinear SOC model through our explicit data-driven approach and utilizing experimental battery measurements. This model underwent training and validation on an in-house stochastic current profile, with corresponding voltage and SOC data sampled at 50 ms intervals. It attained a training RMSE of 2.2×10^{-6} and a validation RMSE of 4.8×10^{-4}. To assess its generalizability, we conducted cross-validation on experimental battery measurements corresponding to the US06 drive cycle, achieving an RMSE of 8.5×10^{-4}. The resulting model is concise, comprising only eight terms selected from the augmented library Θ_a,

including our baseline library Θ_{bl} and some additional nonlinear terms identified via the Monte Carlo Library Search (MCLS). A summary of the findings is presented in Table 5.3. The validation and cross-validation results are illustrated in Fig 5.3a and Fig 5.3b, respectively. We note that comparable performance was achieved by the model developed using idealized simulated data. The predictive performance of this model across the training, validation, and cross-validation tests is also summarized in Table 5.3.

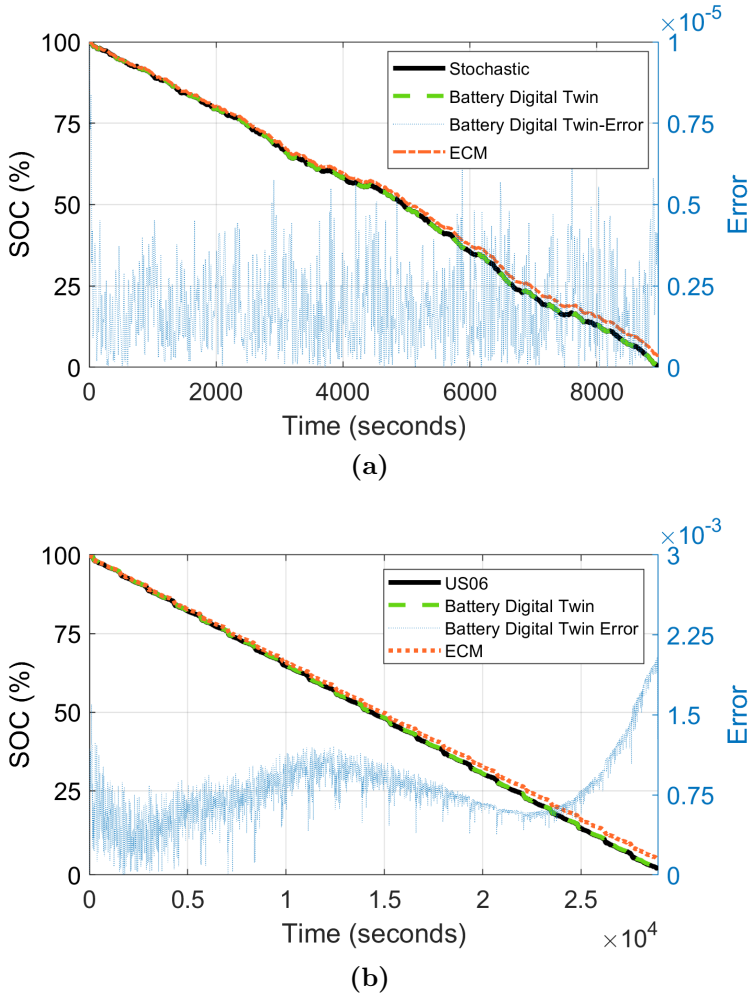

Figure 5.3. Digital Twin Results a) Battery Digital Twin Validation Results: Experimental Stochastic Cycle Data at 25°C and b) Battery Digital Twin Cross-Validation Results: Experimental US06 Cycle Data at 25°C.

We also compared our approach to the commonly used ECM method by fine-tuning its parameters with our experimental training data. The performance of both methods is presented in Fig. 5.3a and Fig. 5.3b. Fig. 5.3a demonstrates that the ECM performs poorly compared to our model, especially in the low SOC region. The ECM's prediction error (RMSE) is 2.4×10^{-2}, significantly higher than the 2.2×10^{-6} achieved by our method. Moreover, the ECM's accuracy is consistent across new data as shown in Fig. 5.3b, where a generalization RMSE of 2.5×10^{-2} was achieved. These results suggest that our approach offers significant improvements over the ECM method, particularly at low SOC levels where the ECM struggles. Additionally, our model simplifies implementation by eliminating the need for multiple sets of coefficients to cover the full range of SOC, a common requirement for the ECM.

Lastly, we developed a re-calibration approach (see §3.1.4) to re-optimize the model coefficients for maintaining predictive performance (RMSE) across diverse temperatures, ranging from -20°C to 40°C. This was done while keeping the optimal model structure (feature library, refer to Table 5.3) with minimal complexity. The optimization process aimed to minimize an RMSE-based cost function (3.10) by recalibrating each model coefficient using data from our stochastic cycle conducted at discrete temperature intervals. The optimal trend of model coefficients is shown in Fig. 5.4a. These trends were achieved through multiple iterations of the optimization routine, starting with a coarse resolution of the coefficient search space, allowing a maximum deviation of 10% from the optimal coefficients of the base model developed at the standard operating temperature of 25°C. In each iteration, the search space was refined, increasing the resolution as we focused on areas of good performance. Coefficients C1 through C8 correspond to each of the eight model terms, encompassing SOC, V, and I. It's worth mentioning that

coefficients related to V (C7 and C8) exhibit the most significant temperature dependency, whereas the SOC and I terms (C1 to C6) undergo minimal or no alterations with temperature variations. The temperature dependency observed in the V terms can be attributed to fluctuations in the battery's voltage response influenced by shifts in operating temperatures [76, 195]. The ultimate model maintained high accuracy across the entire operating range, featuring our optimal feature library and a lookup table of temperature-dependent coefficients optimized for each discrete temperature condition. This includes temperatures from -20°C to 40°C and SOC values from 0% to 100%, with an average RMSE of 1.1×10^{-3}. The RMSE results for each discrete temperature are shown in Fig. 5.4b.

5.5 Summary and Conclusion

This study presents a novel physics-informed battery digital twin (PhITEDD), focusing on accurately predicting State of Charge (SOC) dynamics across various temperatures and SOC conditions. Leveraging battery operando measurements, our digital twin employs an explicit data-driven approach to uncover governing equations for precise SOC forecasting. Our digital twin model is constructed using a reduced-order framework. It comprises a library of candidate terms and coefficients determined through a sparsity-promoting algorithm. We enriched the model's library with explicit physics-informed terms to enhance interpretability and generalizability. We developed a Monte Carlo search strategy to effectively explore further nonlinear terms, enhancing our ability to explore the vast search space and better characterize highly nonlinear behaviors. A hyperparameter autotuning technique was crafted for the regularization optimizer (STRidge) to determine optimal coefficients, striking a balance between model accuracy and complexity. Furthermore, we devised a re-calibration approach to optimize

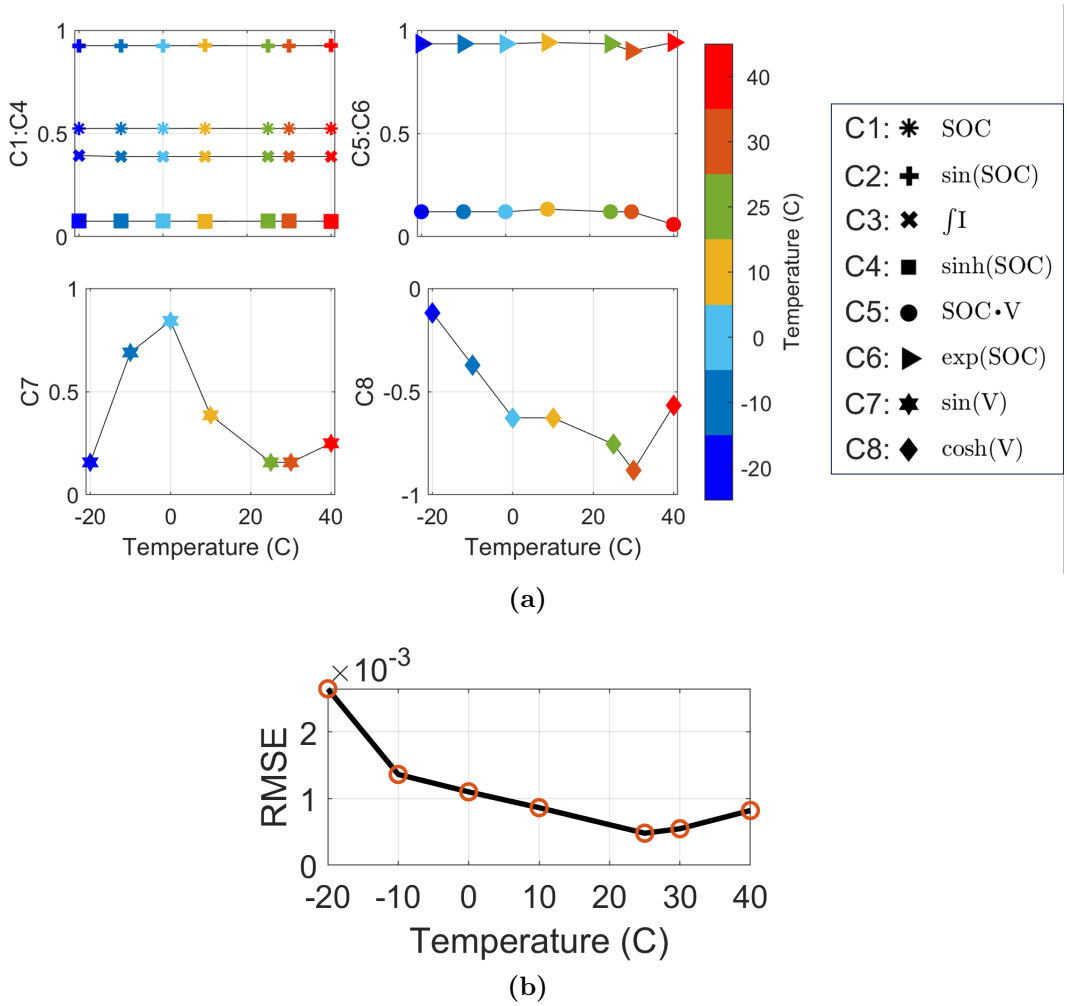

Figure 5.4. PhITEDD Model a) Optimal Trend of Temperature-Dependent Model Coefficients ($\Xi^*_{T_i}$) and b) Predictive Performance Across Temperature Conditions.

model coefficients based on new data, ensuring consistent efficacy across a wide temperature range (-20°C to 40°C) while maintaining minimal complexity. We examined how varying data sampling rates affect the accuracy of data-driven battery models. Subsequently, we optimized the sampling rate and confirmed our findings through pulse relaxation studies. The base model, trained and validated using an in-house stochastic drive cycle at 25°C, demonstrated high accuracy with RMSE scores of 2.2×10^{-6} and 4.8×10^{-4}, respectively, with a parsimonious structure comprising only eight terms. It exhibited strong generalization performance with an RMSE of 8.5×10^{-4} on unseen measurements corresponding to the US06 drive cycle. Overall, the PhITEDD model achieved an average RMSE of 1.1×10^{-3} across the entire operational spectrum, demonstrating its adaptability and the effectiveness of our modeling approach in addressing diverse conditions. Moreover, the nonlinear nature of PhITEDD and its connection to physics provides improvements over ECM-based methods, which show deficiencies at low temperatures and low SOC conditions. In our stochastic and US06 drive cycle experiments (see Fig.5.3a-5.3b), the ECM's performance deteriorated by up to 50% at low SOC levels, while PhITEDD maintained its efficacy. Additionally, it is crucial to emphasize that our modeling approach requires significantly less data than other machine learning techniques, enabling faster training and re-calibration. This advantage facilitates the efficient development of fine-tuned models for individual cells, helping to mitigate errors arising from the inherent inconsistencies within battery packs, which are exacerbated by varying aging progressions. However, it is important to note that the RMSE of multiple battery models will accumulate. Lastly, it is worth noting that the method presented can be tailored to various energy storage setups, including different battery systems,

cell types, and chemistries, and can also serve as a guide for machine learning modeling of complex systems.

6. FAST CHARGING OPTIMIZATION RESULTS

This section presents the optimal charging strategy synthesized with our adaptive learning and optimization method. Our approach employed a full-order electrochemical (DFN) model coupled with a thermal model for capturing the battery's thermal effects, making our electrochemical-thermal-based control law close to the actual battery mechanism.

We explored different charging strategies, including passive changing strategies in §6.1. Our optimal results are presented in §6.2 along with a comparison to the other charging strategies.

6.1 Passive Charging Strategies

Passive charging techniques are model-free methods that charge the battery under preset instructions, as shown in Fig. 6.1. The charging profiles developed with these methods are characterized by their fixed terminal conditions, including current, voltage, or power constraints. However, passive charging algorithms do not consider the feedback of the battery states, which may lead to a shortened battery lifespan.

A common and arguably most widely used charging strategy is constant-current constant-voltage (CC-CV) due to its easy implementation and operation. This algorithm initially charges the battery with a constant current until the voltage reaches a preset upper limit. Then, the voltage is held constant until the current is reduced to a preset minimum value. In this study, we tested the CC-CV charging protocol under different changing rates (C-rates) to fully charge

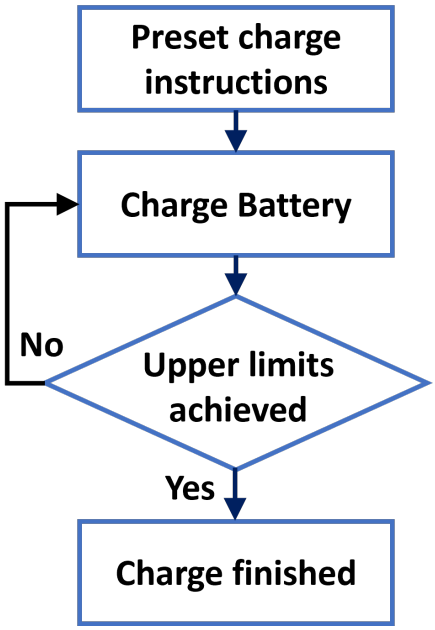

Figure 6.1. Passive Charging Structure

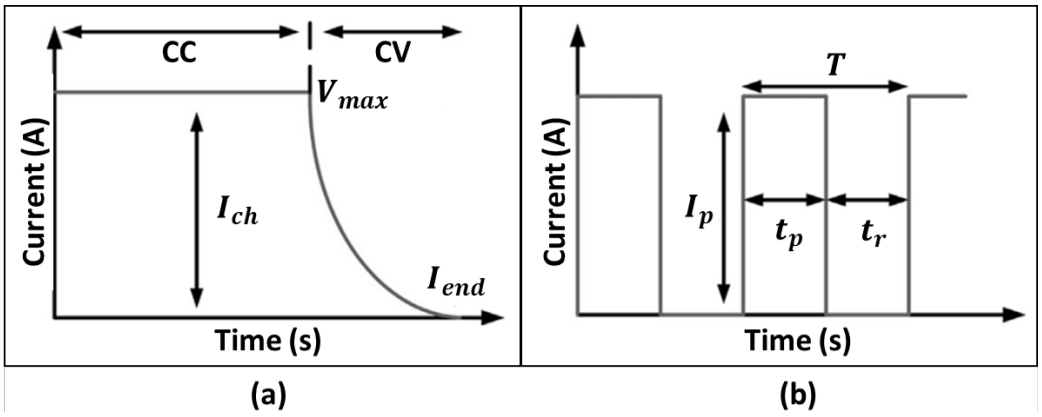

Figure 6.2. Passive Charging Strategies: (a) CC-CV, (b) PPC

a 5Ah 21700 NMC-811 cylindrical cell. First, we tested the conventional 0.3C CC-CV charging strategy following the battery manufacturer's specifications. Next, we tested a fast-charging 2C CC-CV charging protocol, which reduced the charging time to around 4000 seconds from the 12000 seconds needed to fully charge the cell with the 0.3C CC-CV. However, this reduction in time came

Strategy	**Charge Time (s)**	**Max T (C)**	**Max V (V)**
CC-CV: 0.3C	12,500	27	4.20
CC-CV: 2.0C	4,000	64	4.20
Hybrid:	4,000	57	4.22

Table 6.1. Comparison of Charging Strategies

with a substantial rise in the battery's temperature, surpassing the maximum temperature of $63°C$. Moreover, subjecting the battery to elevated temperatures can lead to adverse effects on battery health, such as accelerated electrochemical aging [18, 21]. The CC-CV charging protocols are baselines for comparison against our constrained-based optimal solution, which aims to fulfill fast charging demands while maintaining safe operating conditions by respecting constraints. The corresponding plots and the summarized results are presented in §6.2.

6.2 Optimal Results

Here, we present the optimal charging profile developed with our adaptive optimization approach to maximize the battery SOC within a set charging duration (t_f) while respecting safety constraints. This is achieved by minimizing the square error between the SOC reached during the iteration and the desired SOC (SOC_d) of 100% (fully charged). The optimization objective and operational constraints are defined in (6.1) and summarized in Table 3.2.

$$I^* = \arg\min_I \int_0^{t_f} (SOC(t) - SOC_d)^2 dt \tag{6.1}$$

subject to the constraints:

$$T(t) \leq T_{ub}$$

$$V_{lb} \leq V(t) \leq V_{ub}$$

$$I_{lb} \leq I(t) \leq I_{ub}$$

Our optimized charging strategy comprises a hybrid (mixed continuous-discrete) solution, where continuous refers to the direct simulation of operating modes (e.g., CC, CV, pulse), and discrete refers to a transition between the operating modes. This approach aims to maximize current and subsequently dynamically transition between operating modes to meet constraints. Since the battery has a smaller resistance in the lower SOC range, the highest current is applied as a positive pulse current (PPC), whose waveform parameters (refer to Fig. 6.2) such as peak charging current (I_p), pulse on-time (t_p), relaxation interval time (t_r), and total pulse period (T) are optimized via our adaptive learning and control approach. Pulse charging was implemented, as it can be an efficient and fast charging strategy that, with proper selection of current waveform parameters, can help prevent the side reactions caused by saturation at the particle interface [196]. Following the PPC mode, the solution switches to CV to avoid continuing temperature rise due to the battery's rapidly increasing internal resistance.

Our optimization approach initializes with information (e.g., t_f, I_{ch}, I_{end}, etc.) from the 2C CC-CV profile. It optimizes a set of control points (PPC parameters) to yield a fast charge time while respecting safety constraints, including a maximum voltage of 4.2V and a maximum temperature of 57 $^{\circ}C$ corresponding

to 90% of the maximum surface temperature of $63^{\circ}C$. The optimized charging strategy fully charged the 5Ah 21700 NMC-811 cylindrical cell, 66% faster than the recommended 0.3C CC-CV strategy. It maintained a temperature of $57^{\circ}C$ or lower while the 2C CC-CV strategy experienced higher temperatures, reaching upwards of $64^{\circ}C$. The results are summarized in Table 6.1, while the plots are shown in Fig 6.3.

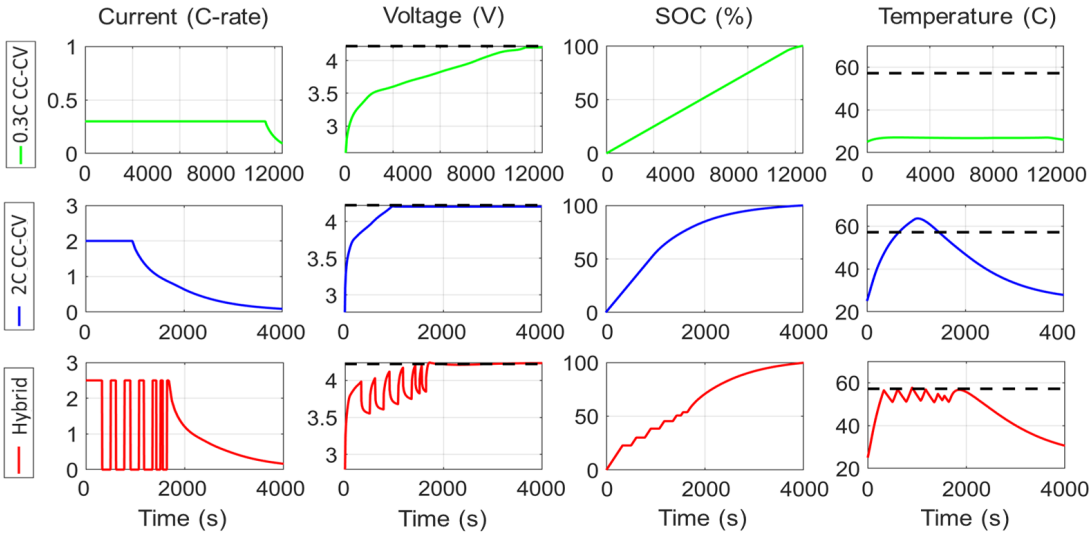

Figure 6.3. Comparison of Charging Strategies

6.3 Conclusions

In this work, we developed a constrained optimal charging strategy that meets fast charging demands and sustains LiBs' safe operation. To avoid subjecting the cell to accelerated aging, we propose optimizing the electrical current for minimum battery charge time while respecting safety constraints, including a maximum cell temperature and a maximum voltage. We used a control strategy to learn the Jacobian of a closed-loop system from input/output data generated by a full-order electrochemical-thermal battery model. Based on the learned dynamics,

we optimized the response. Our optimized charging strategy is comprised of a hybrid (mixed continuous-discrete) solution that fully charges a 5Ah 21700 NMC-811 cylindrical cell, 66% faster than the recommended 0.3C constant-current constant-voltage (CC-CV) strategy. Furthermore, it maintained a temperature of $57°C$ (90% of the $63°C$ maximum temperature) or lower while a comparable 2C CC-CV strategy experienced higher temperatures surpassing $63°C$, which can lead to adverse effects on battery health.

7. SUMMARY AND CONCLUSION

This dissertation makes significant advancements in Li-ion battery modeling and control by introducing a comprehensive, physics-inspired, data-driven framework designed to address the challenges of accurate state prediction and optimal fast charging across diverse operational conditions. The development of PhITEDD, a temperature-dependent battery digital twin, marks a major step forward in creating interpretable, generalizable, and high-accuracy models capable of real-time state-of-charge forecasting. Additionally, the dissertation tackles the critical problem of fast-charging optimization, presenting a robust data-driven control strategy that minimizes charging time while ensuring safety constraints are met, thereby mitigating degradation and enhancing battery longevity.

Key contributions include the advancement of data-driven modeling techniques to develop high-fidelity models of energy storage systems using non-invasive input/output data. These models were developed using a sparsity-promoting optimization algorithm applied to a comprehensive library of candidate terms, which was further enriched with explicit physics-inspired features relevant to the behavior of Lithium-ion batteries. The library incorporated both empirical and theoretical terms that relate to key battery phenomena, such as diffusion processes, intercalation, and electrochemical kinetics. By leveraging the structure of the SINDy (Sparse Identification of Nonlinear Dynamical Systems) framework, the approach ensures that the resulting model captures the essential nonlinear dynamics of the battery while promoting model sparsity. This approach not only

simplifies the model but also enhances its generalizability and interpretability, which are critical for real-world applications.

The integration of physics-informed learning into SINDy marks a significant advancement in the field of control-oriented battery modeling. Unlike traditional methods, which may rely on complex, data-driven approaches or oversimplified models, this novel method benefits from both a connection to the underlying battery physics and data-driven flexibility. The ability to span the entire State-of-Charge (SOC) range, from 0% to 100%, using a single model with a consistent set of coefficients represents a major improvement. Traditional state-of-the-art methods, such as equivalent circuit models (ECMs), typically require multiple sets of coefficients to represent different portions of the SOC range, which introduces complexity and limits model applicability across varying conditions.

To further enhance the modeling framework, a Monte Carlo search algorithm was developed to identify and incorporate additional nonlinear terms, significantly improving the representation of complex system dynamics. This algorithm efficiently navigates high-dimensional feature spaces, enabling the discovery of terms that capture intricate interactions within the system that might otherwise be overlooked.

In addition, a hyperparameter auto-tuning approach was implemented to optimize the balance between model accuracy and complexity. This automated process eliminates the need for manual hyperparameter selection, which can be time-consuming and error-prone, particularly when dealing with high-dimensional systems. By dynamically adjusting the hyperparameters, the method ensures that the model maintains high accuracy while avoiding overfitting, ultimately resulting in a more robust and reliable framework.

To address variability in operating conditions, a recalibration strategy was introduced to update model coefficients using new data. This approach ensures that the model adapts to changes in operating conditions, such as variations in temperature while maintaining simplicity. Specifically, the recalibration process allows the model to perform consistently across a wide temperature range, from $(-20°C$ to $-40°C)$, ensuring accurate predictions and reliable performance even under extreme conditions.

Together, these advancements; Monte Carlo-based exploration of nonlinear terms, hyperparameter auto-tuning, and a robust recalibration mechanism; create a powerful and adaptable modeling framework. This framework is well-suited for applications requiring high precision and flexibility, such as battery management systems in electric vehicles, where maintaining consistent performance across diverse operational scenarios is critical.

The study also explored the impact of varying data sampling rates on the accuracy of data-driven battery models, optimizing the sampling rate and validating the findings through pulse relaxation studies. The methodology was rigorously validated using both simulated and experimental data. Models trained and validated on simulated data from the standard UDDS city driving cycle achieved impressive predictive accuracy, with error values of $(E_t = 5 \times 10^{-7})$ and $(E_v = 1.3 \times 10^{-3})$, respectively. Cross-validation on the unseen US06 highway driving cycle further demonstrated the model's robust generalizability, with an error of $(E_{cv} = 1.4 \times 10^{-3})$.

The model's adaptability was confirmed through experimental data, where the base model, trained and validated using an in-house stochastic drive cycle at 25°C, achieved error values of $(E_t = 2.2 \times 10^{-6})$ and $(E_v = 4.8 \times 10^{-4})$, respectively, using a parsimonious structure of only eight terms. The model also exhibited

strong generalization, with an error of ($E_{cv} = 8.5 \times 10^{-4}$) on unseen real-world US06 drive cycle data.

Overall, the PhITEDD model achieved an average error of (1.1×10^{-3}) across its operational range including temperatures ranging from $-20°C$ to $-40°C$, low (0%) and high (100%) SOC levels, and aggressive currents, demonstrating its adaptability and effectiveness under various conditions. The nonlinear nature and physics-based approach of PhITEDD provide significant improvements over traditional ECM methods, which often struggle under low-temperature and low-SOC conditions. In the stochastic and US06 drive cycle experiments, ECM performance deteriorated by up to 50% at low SOC levels, while PhITEDD maintained high accuracy.

Furthermore, this approach requires significantly less data than other machine learning methods, enabling faster training and recalibration. This efficiency supports the development of fine-tuned models for individual cells, helping mitigate errors caused by inconsistencies within battery packs, particularly those arising from varying aging progressions.

In addition to modeling advancements, this work develops a constrained optimal charging strategy to meet fast-charging demands while ensuring safe operation. The strategy optimizes electrical current to minimize charging time while adhering to safety constraints, such as maximum cell temperature and voltage, to prevent accelerated aging. Using a full-order electrochemical-thermal model, a control strategy was employed to learn the Jacobian of the closed-loop system and optimize the response. The resulting hybrid (mixed continuous-discrete) charging solution achieved a 66% faster charge time for a 5Ah 21700 NMC-811 cylindrical cell compared to the 0.3C constant-current constant-voltage (CC-CV) strategy, while maintaining temperatures below $57°C$, a 10% buffer from

the critical $63°C$ threshold. In contrast, a 2C CC-CV strategy exceeded this limit, posing risks to battery health. Future work will focus on expanding optimization criteria to include minimizing capacity fade and improving efficiency by replacing complex electrochemical models with high-fidelity, data-driven reduced-order models.

This dissertation not only contributes transformative methodologies for Li-ion battery modeling and control but also establishes a foundation for broader applications in dynamic systems modeling. By bridging gaps between data-driven and physics-based approaches, this work offers significant potential to advance battery technologies, ensuring faster, safer, and more sustainable energy storage solutions for critical applications like electric vehicles and renewable energy systems.

Future work will aim to extend our modeling framework PhITEDD, to better capture battery aging, which is crucial for long-term performance and reliability. This extension will focus on (i) incorporating aging mechanisms such as capacity fade, increased internal resistance, and degradation of materials; (ii) using long-term experimental data from accelerated aging tests to refine the model and improve accuracy; and (iii) integrating dynamic aging parameters that adjust based on usage patterns, environmental conditions, and charge/discharge cycles, with real-time updates to maintain prediction accuracy as the battery ages. Furthermore, on the optimal charging front, we aim to expand our optimization criteria to include minimizing damage to the cyclable life of the battery quantified by capacity fade. Also, we plan on improving the efficiency of our optimization approach by substituting the complex electrochemical model with our PhITEDD Battery Digital Twin.

REFERENCES CITED

[1] International Energy Agency. Global ev outlook 2024, 2024. https://www.iea.org/reports/global-ev-outlook-2024 [Accessed: (2024-1-1)].

[2] Gregory L. Plett. *Battery Management Systems: Volume 1: Battery Modeling*. Artech House Power Engineering and Power Electronics. Artech House, Boston London, 2015.

[3] Ming Jiang, Dongjiang Li, Zonghua Li, Zhuo Chen, Qinshan Yan, Fu Lin, Cheng Yu, Bo Jiang, Xuezhe Wei, and Wensheng Yan. Advances in battery state estimation of battery management system in electric vehicles. *Journal of Power Sources*, 612:234781, 2024.

[4] F. Naseri, S. Gil, C. Barbu, E. Cetkin, G. Yarimca, A. C. Jensen, P. G. Larsen, and C. Gomes. Digital twin of electric vehicle battery systems: Comprehensive review of the use cases, requirements, and platforms. *Renewable and Sustainable Energy Reviews*, 179:113280, June 2023.

[5] Prashant Shrivastava, Tey Kok Soon, Mohd Yamani Idna Bin Idris, and Saad Mekhilef. Overview of model-based online state-of-charge estimation using Kalman filter family for lithium-ion batteries. *Renewable and Sustainable Energy Reviews*, 113:109233, October 2019.

[6] Gregory L. Plett. *Battery Management Systems: Volume 2: Equivalent-Circuit Methods*. Artech House Power Engineering Series. Artech house, Boston, 2016.

[7] Tahmineh Raoofi and Melih Yildiz. Comprehensive review of battery state estimation strategies using machine learning for battery Management Systems of Aircraft Propulsion Batteries. *Journal of Energy Storage*, 59:106486, March 2023.

[8] Adam Smiley and Gregory L. Plett. An adaptive physics-based reduced-order model of an aged lithium-ion cell, selected using an interacting multiple-model Kalman filter. *Journal of Energy Storage*, 19:120–134, October 2018.

[9] Domenico Natella, Simona Onori, and Francesco Vasca. A Co-Estimation Framework for State of Charge and Parameters of Lithium-Ion Battery With Robustness to Aging and Usage Conditions. *IEEE Trans. Ind. Electron.*, 70(6):5760–5770, June 2023.

[10] Renato Rodriguez, Omidreza Ahmadzadeh, Yan Wang, and Damoon Soudbakhsh. Data-Driven Discovery of Lithium-Ion Battery State of Charge Dynamics. *Journal of Dynamic Systems, Measurement, and Control*, pages 1–13, 2024.

[11] Zhongbao Wei, Changfu Zou, Feng Leng, Boon Hee Soong, and King-Jet Tseng. Online Model Identification and State-of-Charge Estimate for Lithium-Ion Battery With a Recursive Total Least Squares-Based Observer. *IEEE Transactions on Industrial Electronics*, 65(2):1336–1346, February 2018.

[12] Li Zhang, Kang Li, Dajun Du, Chunbo Zhu, and Min Zheng. A sparse least squares support vector machine used for soc estimation of li-ion batteries. *IFAC-PapersOnLine*, 52(11):256–261, 2019.

[13] Renato Rodriguez, Omidreza Ahmadzadeh, Yan Wang, and Damoon Soudbakhsh. Discovering governing equations of li-ion batteries pertaining state of charge using input-output data. In *2023 American Control Conference (ACC)*, pages 3081–3086. IEEE, 2023.

[14] Omidreza Ahmadzadeh, Renato Rodriguez, and Damoon Soudbakhsh. Modeling of Li-ion batteries for real-time analysis and control: A data-driven approach. In *ACC'2022*, pages 392–397, 2022.

[15] Qizhe Lin, Xiaoqi Li, Bicheng Tu, Junwei Cao, Ming Zhang, and Jiawei Xiang. Stable and Accurate Estimation of SOC Using eXogenous Kalman Filter for Lithium-Ion Batteries. *Sensors*, 23(1):467, January 2023.

[16] Saad El Fallah, Jaouad Kharbach, Zakia Hammouch, Abdellah Rezzouk, and Mohammed Ouazzani Jamil. State of charge estimation of an electric vehicle's battery using Deep Neural Networks: Simulation and experimental results. *Journal of Energy Storage*, 62:106904, June 2023.

[17] Daniel Juarez Robles. *Degradation-Safety Analytics in Energy Storage*. PhD thesis, Purdue University Graduate School, 2019.

[18] Marc D Berliner, Daniel A Cogswell, Martin Z Bazant, and Richard D Braatz. A mixed continuous-discrete approach to fast charging of li-ion batteries while maximizing lifetime. *IFAC-PapersOnLine*, 55(30):305–310, 2022.

[19] Carlos Pastor-Fernández, Kotub Uddin, Gael H. Chouchelamane, W. Dhammika Widanage, and James Marco. A Comparison between Electrochemical Impedance Spectroscopy and Incremental Capacity-Differential Voltage as Li-ion Diagnostic Techniques to Identify and Quantify the Effects of Degradation Modes within Battery Management Systems. *Journal of Power Sources*, 360:301–318, August 2017.

[20] Wenwen Wang, Jun Wang, Jinpeng Tian, Jiahuan Lu, and Rui Xiong. Application of Digital Twin in Smart Battery Management Systems. *Chinese Journal of Mechanical Engineering*, 34(1):57, December 2021.

[21] Anna Tomaszewska, Zhengyu Chu, Xuning Feng, Simon O'kane, Xinhua Liu, Jingyi Chen, Chenzhen Ji, Elizabeth Endler, Ruihe Li, and Lishuo Liu. Lithium-ion battery fast charging: A review. *ETransportation*, 1:100011, 2019.

[22] Donald J. Docimo. Estimation and balancing of multi-state differences between lithium-ion cells within a battery pack. *Journal of Energy Storage*, 50:104264, 2022.

[23] Yizhao Gao, Xi Zhang, Qiyu Cheng, Bangjun Guo, and Jun Yang. Classification and review of the charging strategies for commercial lithium-ion batteries. *Ieee Access*, 7:43511–43524, 2019.

[24] Benedikt Rzepka, Simon Bischof, and Thomas Blank. Implementing an extended kalman filter for soc estimation of a li-ion battery with hysteresis: A step-by-step guide. *Energies*, 14(13):3733, 2021.

[25] Chang-Hui Chen, Ferran Brosa Planella, Kieran O'regan, Dominika Gastol, W Dhammika Widanage, and Emma Kendrick. Development of experimental techniques for parameterization of multi-scale lithium-ion battery models. *Journal of The Electrochemical Society*, 167(8):080534, 2020.

[26] Caiping Zhang, Jiuchun Jiang, Linjing Zhang, Sijia Liu, Leyi Wang, and Poh Chiang Loh. A generalized soc-ocv model for lithium-ion batteries and the soc estimation for lnmco battery. *Energies*, 9(11):900, 2016.

[27] Joel C Forman, Scott J Moura, Jeffrey L Stein, and Hosam K Fathy. Genetic identification and fisher identifiability analysis of the doyle–fuller–newman model from experimental cycling of a lifepo4 cell. *Journal of Power Sources*, 210:263–275, 2012.

[28] Robert Timms, Scott G Marquis, Valentin Sulzer, Colin P Please, and S Jonathan Chapman. Asymptotic reduction of a lithium-ion pouch cell model. *SIAM Journal on Applied Mathematics*, 81(3):765–788, 2021.

[29] Amardeep Sidhu, Afshin Izadian, and Sohel Anwar. Adaptive nonlinear model-based fault diagnosis of li-ion batteries. *IEEE Transactions on Industrial Electronics*, 62(2):1002–1011, 2014.

[30] Quan-Qing Yu, Rui Xiong, Le-Yi Wang, and Cheng Lin. A comparative study on open circuit voltage models for lithium-ion batteries. *Chinese Journal of Mechanical Engineering*, 31(1):1–8, 2018.

[31] Carlos Vidal, Pawel Malysz, Mina Naguib, Ali Emadi, and Phillip J Kollmeyer. Estimating battery state of charge using recurrent and non-recurrent neural networks. *Journal of Energy Storage*, 47:103660, 2022.

[32] Simon EJ O'Kane, Weilong Ai, Ganesh Madabattula, Diego Alonso-Alvarez, Robert Timms, Valentin Sulzer, Jacqueline Sophie Edge, Billy Wu, Gregory J Offer, and Monica Marinescu. Lithium-ion battery degradation: how to model it. *Physical Chemistry Chemical Physics*, 24(13):7909–7922, 2022.

[33] Valentin Sulzer, Scott G Marquis, Robert Timms, Martin Robinson, and S Jon Chapman. Python battery mathematical modelling (pybamm). *Journal of Open Research Software*, 9(1), 2021.

[34] Alexander Cho, Stephen Nah, Greyson Sapio, Daniel Vail, Patrick Wang, and Thomas Hodson. Analysis of lithium-ion battery failure and pybamm's viability in simulating them. In *2020 IEEE MIT Undergraduate Research Technology Conference (URTC)*, pages 1–4. IEEE, 2020.

[35] Daniel Luder, Priscilla Caliandro, and Andrea Vezzini. Enhanced physics-based models for state estimation of li-ion batteries. In *Proceedings of the Comsol Conference*, 2020.

[36] Wenlin Zhang, Ryan Ahmed, and Saeid Habibi. The effects of test profile on lithium-ion battery equivalent-circuit model parameterization accuracy. In *2022 IEEE Transportation Electrification Conference & Expo (ITEC)*, pages 119–124. IEEE, 2022.

[37] Yasser Ghoulam, Tedjani Mesbahi, Peter Wilson, Sylvain Durand, Andrew Lewis, Christophe Lallement, and Christopher Vagg. Lithium-ion battery parameter identification for hybrid and electric vehicles using drive cycle data. *Energies*, 15(11):4005, 2022.

[38] Mohsen Derakhshan and Damoon Soudbakhsh. Temperature-dependent time constants of li-ion batteries. *IEEE Control Systems Letters*, 6:2012–2017, 2021.

[39] Yuntian Chen, Yingtao Luo, Qiang Liu, Hao Xu, and Dongxiao Zhang. Any equation is a forest: Symbolic genetic algorithm for discovering open-form partial differential equations (SGA-PDE). *arXiv preprint arXiv:2106.11927*, 2021.

[40] Saehong Park, Dylan Kato, Zach Gima, Reinhardt Klein, and Scott Moura. Optimal experimental design for parameterization of an electrochemical lithium-ion battery model. *J. Electrochem. Soc.*, 165(7):A1309, 2018.

[41] Saehong Park, Dong Zhang, Reinhardt Klein, and Scott Moura. Estimation of cyclable lithium for li-ion battery state-of-health monitoring. In *2021 American Control Conference (ACC)*, pages 3094–3101. IEEE, 2021.

[42] Scott G Marquis, Valentin Sulzer, Robert Timms, Colin P Please, and S Jon Chapman. An asymptotic derivation of a single particle model with electrolyte. *Journal of The Electrochemical Society*, 166(15):A3693, 2019.

[43] Scott G Marquis, Robert Timms, Valentin Sulzer, Colin P Please, and S Jon Chapman. A suite of reduced-order models of a single-layer lithium-ion pouch cell. *Journal of The Electrochemical Society*, 167(14):140513, 2020.

[44] Jan Kleiner, Magdalena Stuckenberger, Lidiya Komsiyska, and Christian Endisch. Advanced monitoring and prediction of the thermal state of intelligent battery cells in electric vehicles by physics-based and data-driven modeling. *Batteries*, 7(2):31, 2021.

[45] Dong Zhang, Satadru Dey, Luis D Couto, and Scott J Moura. Battery adaptive observer for a single-particle model with intercalation-induced stress. *IEEE Trans Control Syst Technol*, 28(4):1363–1377, 2019.

[46] Hao Tu, Scott Moura, and Huazhen Fang. Integrating electrochemical modeling with machine learning for lithium-ion batteries. In *2021 American Control Conference (ACC)*, pages 4401–4407. IEEE, 2021.

[47] Hao Tu, Scott Moura, Yebin Wang, and Huazhen Fang. Integrating physics-based modeling with machine learning for lithium-ion batteries. *Applied Energy*, 329:120289, 2023.

[48] Yinjiao Xing, Wei He, Michael Pecht, and Kwok Leung Tsui. State of charge estimation of lithium-ion batteries using the open-circuit voltage at various ambient temperatures. *Applied Energy*, 113:106–115, 2014.

[49] Caihao Weng, Jing Sun, and Huei Peng. A unified open-circuit-voltage model of lithium-ion batteries for state-of-charge estimation and state-of-health monitoring. *Journal of power Sources*, 258:228–237, 2014.

[50] Fangdan Zheng, Yinjiao Xing, Jiuchun Jiang, Bingxiang Sun, Jonghoon Kim, and Michael Pecht. Influence of different open circuit voltage tests on state of charge online estimation for lithium-ion batteries. *Applied energy*, 183:513–525, 2016.

[51] Kong Soon Ng, Chin-Sien Moo, Yi-Ping Chen, and Yao-Ching Hsieh. Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries. *Applied energy*, 86(9):1506–1511, 2009.

[52] Asep Nugroho, Estiko Rijanto, F Danang Wijaya, and Prapto Nugroho. Battery state of charge estimation by using a combination of coulomb counting and dynamic model with adjusted gain. In *2015 International Conference on Sustainable Energy Engineering and Application (ICSEEA)*, pages 54–58. IEEE, 2015.

[53] Edi Leksono, Irsyad Nashirul Haq, Muhammad Iqbal, FX Nugroho Soelami, and IGN Merthayasa. State of charge (soc) estimation on lifepo 4 battery module using coulomb counting methods with modified peukert. In *2013 Joint International Conference on Rural Information & Communication Technology and Electric-Vehicle Technology (rICT & ICeV-T)*, pages 1–4. IEEE, 2013.

[54] Ruxiu Zhao, Phillip J Kollmeyer, Robert D Lorenz, and Thomas M Jahns. A compact methodology via a recurrent neural network for accurate equivalent circuit type modeling of lithium-ion batteries. *IEEE Transactions on Industry Applications*, 55(2):1922–1931, 2018.

[55] Manh-Kien Tran, Anosh Mevawala, Satyam Panchal, Kaamran Raahemifar, Michael Fowler, and Roydon Fraser. Effect of integrating the hysteresis component to the equivalent circuit model of lithium-ion battery for dynamic and non-dynamic applications. *Journal of Energy Storage*, 32:101785, 2020.

[56] Ephrem Chemali, Phillip J Kollmeyer, Matthias Preindl, and Ali Emadi. State-of-charge estimation of Li-ion batteries using deep neural networks: A machine learning approach. *J. Power Sources*, 400:242–255, 2018.

[57] Carlos Vidal, Pawel Malysz, Mina Naguib, Ali Emadi, and Phillip J Kollmeyer. Estimating battery state of charge using recurrent and non-recurrent neural networks. *Journal of Energy Storage*, 47:103660, 2022.

[58] Zahra Nozarijouybari and Hosam K. Fathy. Machine learning for battery systems applications: Progress, challenges, and opportunities. *Journal of Power Sources*, 601:234272, May 2024.

[59] Md Sazzad Hosen, Joris Jaguemont, Joeri Van Mierlo, and Maitane Berecibar. Battery lifetime prediction and performance assessment of different modeling approaches. *Iscience*, 24(2), 2021.

[60] Hao Tu, Scott Moura, and Huazhen Fang. Integrating Electrochemical Modeling with Machine Learning for Lithium-Ion Batteries. In *2021 American Control Conference (ACC)*, pages 4401–4407, May 2021.

[61] Saehong Park, Dong Zhang, Reinhardt Klein, and Scott Moura. Estimation of cyclable lithium for li-ion battery state-of-health monitoring. In *2021 American Control Conference (ACC)*, pages 3094–3101. IEEE, 2021.

[62] Marc Doyle. Modeling of Galvanostatic Charge and Discharge of the Lithium/Polymer/Insertion Cell. *Journal of The Electrochemical Society*, 140(6):1526, 1993.

[63] Karen E Thomas, John Newman, and Robert M Darling. Mathematical modeling of lithium batteries. In *Advances in Lithium-Ion Batteries*, pages 345–392. Springer, 2002.

[64] Madeleine Ecker, Thi Kim Dung Tran, Philipp Dechent, Stefan Käbitz, Alexander Warnecke, and Dirk Uwe Sauer. Parameterization of a Physico-Chemical Model of a Lithium-Ion Battery: I. Determination of Parameters. *J. Electrochem. Soc.*, 162(9):A1836–A1848, 2015.

[65] Gregory L Plett. *Battery management systems, Volume I: Battery modeling*. Artech House, 2015.

[66] Omidreza Ahmadzadeh, Renato Rodriguez, Yan Wang, and Damoon Soudbakhsh. A Physics-Inspired Machine Learning Nonlinear Model of Li-ion Batteries. In *ACC'2023*, pages 3087–3092, 2023-05.

[67] Shunli Wang, Carlos Fernandez, Yu Chunmei, Fan Yongcun, Cao Wen, Daniel-Ioan Stroe, and Zonghai Chen. *Battery System Modeling*. Elsevier, 2021.

[68] Lei Pei, Rengui Lu, and Chunbo Zhu. Relaxation model of the open-circuit voltage for state-of-charge estimation in lithium-ion batteries. *IET Electrical Systems in Transportation*, 3(4):112–117, 2013.

[69] Bin Yao, Yongxiang Cai, Wei Liu, Yang Wang, Xin Chen, Qiangqiang Liao, Zaiguo Fu, and Zhiyuan Cheng. State-of-charge estimation for lithium-ion batteries based on modified unscented Kalman filter using improved parameter identification. *International Journal of Electrochemical Science*, 19(5):100574, May 2024.

[70] Vikalp Vikalp, Vikalp Jha, Muditha Abeysekera, Nicholas Jenkins, and Jianzhong Wu. Battery digital twin: State of Charge and State of health estimation of LTO battery storage. *Energy Proceedings*, 45, 2024.

[71] Jiamiao Xie, Xingyu Wei, Xiqiao Bo, Peng Zhang, Pengyun Chen, Wenqian Hao, and Meini Yuan. State of charge estimation of lithium-ion battery based on extended Kalman filter algorithm. *Frontiers in Energy Research*, 11, May 2023.

[72] Aihua Tang, Yukun Huang, Shangmei Liu, Quanqing Yu, Weixiang Shen, and Rui Xiong. A novel lithium-ion battery state of charge estimation method based on the fusion of neural network and equivalent circuit models. *Applied Energy*, 348:121578, October 2023.

[73] Shaosen Su, Wei Li, Jianhui Mou, Akhil Garg, Liang Gao, and Jie Liu. A Hybrid Battery Equivalent Circuit Model, Deep Learning, and Transfer Learning for Battery State Monitoring. *IEEE Transactions on Transportation Electrification*, 9(1):1113–1127, March 2023.

[74] Leo Wildfeuer, Philipp Gieler, and Alexander Karger. Combining the Distribution of Relaxation Times from EIS and Time-Domain Data for Parameterizing Equivalent Circuit Models of Lithium-Ion Batteries. *Batteries*, 7(3):52, August 2021.

[75] Xinyou Lin, Yunliang Tang, Jing Ren, and Yimin Wei. State of charge estimation with the adaptive unscented Kalman filter based on an accurate equivalent circuit model. *Journal of Energy Storage*, 41:102840, September 2021.

[76] Cong Jiang, Shunli Wang, Bin Wu, Carlos Fernandez, Xin Xiong, and James Coffie-Ken. A state-of-charge estimation method of the power lithium-ion battery in complex conditions based on adaptive square root extended Kalman filter. *Energy*, 219:119603, 2021.

[77] Yongliang Zheng, Feng He, and Wenliang Wang. A Method to Identify Lithium Battery Parameters and Estimate SOC Based on Different Temperatures and Driving Conditions. *Electronics*, 8(12):1391, November 2019.

[78] Xianzhi Gong. *Modeling of Lithium-Ion Battery Considering Temperature and Aging Uncertainties*. PhD thesis, University of Michigan, 2016.

[79] Donal P. Finegan, Juner Zhu, Xuning Feng, Matt Keyser, Marcus Ulmefors, Wei Li, Martin Z. Bazant, and Samuel J. Cooper. The Application of Data-Driven Methods and Physics-Based Learning for Improving Battery Safety. *Joule*, 5(NREL/JA-5700-76901), December 2020.

[80] Yihuan Li, Kang Li, Xuan Liu, Yanxia Wang, and Li Zhang. Lithium-ion battery capacity estimation—A pruned convolutional neural network approach assisted with transfer learning. *Applied Energy*, 285:116410, 2021.

[81] Hao Tu, Scott Moura, Yebin Wang, and Huazhen Fang. Integrating physics-based modeling with machine learning for lithium-ion batteries. *Applied Energy*, 329:120289, 2023.

[82] Devendrasinh Darbar and Indranil Bhattacharya. Application of Machine Learning in Battery: State of Charge Estimation Using Feed Forward Neural Network for Sodium-Ion Battery. *Electrochem*, 3(1):42–57, January 2022.

[83] Junxiong Chen, Yu Zhang, Wenjiang Li, Weisong Cheng, and Qiao Zhu. State of charge estimation for lithium-ion batteries using gated recurrent unit recurrent neural network and adaptive Kalman filter. *Journal of Energy Storage*, 55:105396, 2022.

[84] Zhenhua Cui, Licheng Wang, Qiang Li, and Kai Wang. A comprehensive review on the state of charge estimation for lithium-ion battery based on neural network. *International Journal of Energy Research*, 46(5):5423–5440, April 2022.

[85] Ala A. Hussein. Capacity fade estimation in electric vehicle li-ion batteries using artificial neural networks. *IEEE Transactions on Industry Applications*, 51(3):2321–2330, 2014.

[86] Fen Zhao, Yinguo Li, Xinheng Wang, Ling Bai, and Tailin Liu. Lithium-ion batteries state of charge prediction of electric vehicles using RNNs-CNNs neural networks. *Ieee Access*, 8:98168–98180, 2020.

[87] Xinyuan Fan, Weige Zhang, Caiping Zhang, Anci Chen, and Fulai An. SOC estimation of Li-ion battery using convolutional neural network with U-Net architecture. *Energy*, 256:124612, 2022.

[88] Xinyou Lin, Hao Huang, Xinhao Xu, and Liping Xie. Dynamic programming solutions extracted SOC-trajectory online learning generation algorithm based approximate global optimization control strategy for a fuel cell hybrid electric vehicle. *Energy*, page 130728, 2024.

[89] Pei Wang, Xue Dan, and Yong Yang. A multi-scale fusion prediction method for lithium-ion battery capacity based on ensemble empirical mode decomposition and nonlinear autoregressive neural networks. *International Journal of Distributed Sensor Networks*, 15(3):1550147719839637, March 2019.

[90] Caihao Weng, Yujia Cui, Jing Sun, and Huei Peng. On-board state of health monitoring of lithium-ion batteries using incremental capacity analysis with support vector regression. *Journal of Power Sources*, 235:36–44, 2013.

[91] Jinhao Meng, Guangzhao Luo, and Fei Gao. Lithium polymer battery state-of-charge estimation based on adaptive unscented Kalman filter and support vector machine. *IEEE Transactions on Power Electronics*, 31(3):2226–2238, 2015.

[92] Xiaosong Hu and Fengchun Sun. Fuzzy clustering based multi-model support vector regression state of charge estimator for lithium-ion battery of electric vehicle. In *2009 International Conference on Intelligent Human-Machine Systems and Cybernetics*, volume 1, pages 392–396. IEEE, 2009.

[93] Xiaosong Hu, Shengbo Eben Li, and Yalian Yang. Advanced Machine Learning Approach for Lithium-Ion Battery State Estimation in Electric Vehicles. *IEEE Transactions on Transportation Electrification*, 2(2):140–149, June 2016.

[94] Mohammad Charkhgard and Mohammad Farrokhi. State-of-charge estimation for lithium-ion batteries using neural networks and EKF. *IEEE transactions on industrial electronics*, 57(12):4178–4187, 2010.

[95] Elad Hoffer, Itay Hubara, and Daniel Soudry. Train longer, generalize better: Closing the generalization gap in large batch training of neural networks. *Advances in neural information processing systems*, 30, 2017.

[96] Yiding Jiang, Dilip Krishnan, Hossein Mobahi, and Samy Bengio. Predicting the Generalization Gap in Deep Networks with Margin Distributions, June 2019.

[97] Muratahan Aykol, Chirranjeevi Balaji Gopal, Abraham Anapolsky, Patrick K. Herring, Bruis van Vlijmen, Marc D. Berliner, Martin Z. Bazant, Richard D. Braatz, William C. Chueh, and Brian D. Storey. Perspective—combining physics and machine learning to predict battery lifetime. *Journal of The Electrochemical Society*, 168(3):030525, 2021.

[98] Jennifer Brucker, Wolfgang G. Bessler, and Rainer Gasper. Grey-box modelling of lithium-ion batteries using neural ordinary differential equations. *Energy Informatics*, 4(S3):15, September 2021.

[99] Chenyu Xue, Bo Jiang, Jiangong Zhu, Xuezhe Wei, and Haifeng Dai. An enhanced single-particle model using a physics-informed neural network considering electrolyte dynamics for lithium-ion batteries. *Batteries*, 9(10):511, 2023.

[100] Chao Hu, Byeng D. Youn, and Jaesik Chung. A multiscale framework with extended Kalman filter for lithium-ion battery SOC and capacity estimation. *Applied Energy*, 92:694–704, 2012.

[101] Saeed Sepasi, Reza Ghorbani, and Bor Yann Liaw. Improved extended Kalman filter for state of charge estimation of battery pack. *Journal of Power Sources*, 255:368–376, 2014.

[102] Fengchun Sun, Xiaosong Hu, Yuan Zou, and Siguang Li. Adaptive unscented Kalman filtering for state of charge estimation of a lithium-ion battery for electric vehicles. *Energy*, 36(5):3531–3540, 2011.

[103] Zhenhua Cui, Le Kang, Liwei Li, Licheng Wang, and Kai Wang. A combined state-of-charge estimation method for lithium-ion battery using an improved BGRU network and UKF. *Energy*, 259:124933, 2022.

[104] Jiahao Li, Joaquin Klee Barillas, Clemens Guenther, and Michael A. Danzer. A comparative study of state of charge estimation algorithms for LiFePO4 batteries used in electric vehicles. *Journal of power sources*, 230:244–250, 2013.

[105] Mehdi Gholizadeh and Farzad R. Salmasi. Estimation of state of charge, unknown nonlinearities, and state of health of a lithium-ion battery based on a comprehensive unobservable model. *IEEE Transactions on Industrial Electronics*, 61(3):1335–1344, 2013.

[106] Xiaopeng Chen, Weixiang Shen, Zhenwei Cao, and Ajay Kapoor. A novel approach for state of charge estimation based on adaptive switching gain sliding mode observer in electric vehicles. *Journal of Power Sources*, 246:667–678, 2014.

[107] Wladislaw Waag, Christian Fleischer, and Dirk Uwe Sauer. Critical review of the methods for monitoring of lithium-ion batteries in electric and hybrid vehicles. *Journal of Power Sources*, 258:321–339, 2014.

[108] Lezhang Liu, Ziqiang Chen, Caisheng Wang, Feng Lin, and Hongbin Wang. Integrated system identification and state-of-charge estimation of battery systems. *IEEE Transactions on Energy conversion*, 28(1):12–23, 2012.

[109] Yong-Min Jeong, Yong-Ki Cho, Jung-Hoon Ahn, Seung-Hee Ryu, and Byoung-Kuk Lee. Enhanced coulomb counting method with adaptive soc reset time for estimating ocv. In *2014 IEEE Energy Conversion Congress and Exposition (ECCE)*, pages 1313–1318. IEEE, 2014.

[110] Ines Baccouche, Sabeur Jemmali, Asma Mlayah, Bilal Manai, and Najoua Essoukri Ben Amara. Implementation of an improved coulomb-counting algorithm based on a piecewise soc-ocv relationship for soc estimation of li-ionbattery. *arXiv preprint arXiv:1803.10654*, 2018.

[111] Alejandro Gismero, Erik Schaltz, and Daniel-Ioan Stroe. Recursive state of charge and state of health estimation method for lithium-ion batteries based on coulomb counting and open circuit voltage. *Energies*, 13(7):1811, 2020.

[112] Damoon Soudbakhsh, Anuradha M Annaswamy, Yan Wang, Steven L Brunton, Joseph Gaudio, Heather Hussain, Draguna Vrabie, Jan Drgona, and Dimitar Filev. Data-Driven Control: Theory and Applications. In *2023 American Control Conference (ACC)*, pages 1922–1939, 2023.

[113] Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *PNAS*, 113(15):3932–3937, 2016.

[114] Omidreza Ahmadzadeh, Yan Wang, and Damoon Soudbakhsh. A data-driven framework for learning governing equations of Li-ion batteries and co-estimating voltage and state-of-charge. *Journal of Energy Storage*, 84:110743, April 2024.

[115] Urban Fasel, Eurika Kaiser, J. Nathan Kutz, Bingni W. Brunton, and Steven L. Brunton. SINDy with Control: A Tutorial. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pages 16–21, December 2021.

[116] Zhanhua Ma, Sunil Ahuja, and Clarence W Rowley. Reduced-order models for control of fluids using the eigensystem realization algorithm. *Theoretical and Computational Fluid Dynamics*, 25(1):233–247, 2011.

[117] Daniel Boley. Computing the kalman decomposition: An optimal method. *IEEE transactions on automatic control*, 29(1):51–53, 1984.

[118] Jer-Nan Juang and Richard S Pappa. Effects of noise on modal parameters identified by the eigensystem realization algorithm. *Journal of Guidance, Control, and Dynamics*, 9(3):294–303, 1986.

[119] Bruce Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE transactions on automatic control*, 26(1):17–32, 1981.

[120] Sanjay Lall, Jerrold E Marsden, and Sonja Glavaški. A subspace approach to balanced truncation for model reduction of nonlinear control systems. *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, 12(6):519–535, 2002.

[121] Björn Liljegren-Sailer. A time domain a posteriori error bound for balancing-related model order reduction. *arXiv preprint arXiv:2301.01052*, 2023.

[122] Clarence W Rowley and Scott TM Dawson. Model reduction for flow analysis and control. *Annu. Rev. Fluid Mech*, 49(1):387–417, 2017.

[123] Y Tamura, S Suganuma, H Kikuchi, and K Hibi. Proper orthogonal decomposition of random wind pressure field. *Journal of Fluids and Structures*, 13(7-8):1069–1095, 1999.

[124] Julien Weiss. A tutorial on the proper orthogonal decomposition. In *AIAA aviation 2019 forum*, page 3333, 2019.

[125] Tony F Chan. An improved algorithm for computing the singular value decomposition. *ACM Transactions on Mathematical Software (TOMS)*, 8(1):72–83, 1982.

[126] H Andrews and CLIII Patterson. Singular value decomposition (svd) image coding. *IEEE transactions on Communications*, 24(4):425–432, 1976.

[127] Michael E Wall, Andreas Rechtsteiner, and Luis M Rocha. Singular value decomposition and principal component analysis. In *A practical approach to microarray data analysis*, pages 91–109. Springer, 2003.

[128] Jer-Nan Juang and Richard S Pappa. An eigensystem realization algorithm for modal parameter identification and model reduction. *Journal of guidance, control, and dynamics*, 8(5):620–627, 1985.

[129] Juan M Caicedo. Practical guidelines for the natural excitation technique (next) and the eigensystem realization algorithm (era) for modal identification using ambient vibration. *Experimental Techniques*, 35(4):52–58, 2011.

[130] Jer-Nan Juang, Minh Phan, Lucas G Horta, and Richard W Longman. Identification of observer/kalman filter markov parameters-theory and experiments. *Journal of Guidance, Control, and Dynamics*, 16(2):320–329, 1993.

[131] MINH Phan, Lucas G Horta, J-N Juang, and Richard W Longman. Improvement of observer/kalman filter identification (okid) by residual whitening. *Journal of Vibration and Acoustics*, 1995.

[132] Francesco Vicario. *OKID as a general approach to linear and bilinear system identification*. Columbia University, 2014.

[133] Joshua L Proctor, Steven L Brunton, and J Nathan Kutz. Dynamic mode decomposition with control. *SIAM Journal on Applied Dynamical Systems*, 15(1):142–161, 2016.

[134] Clarence W Rowley, Tim Colonius, and Richard M Murray. Model reduction for compressible flows using pod and galerkin projection. *Physica D: Nonlinear Phenomena*, 189(1-2):115–129, 2004.

[135] Karen Willcox and Jaime Peraire. Balanced model reduction via the proper orthogonal decomposition. *AIAA journal*, 40(11):2323–2330, 2002.

[136] Renato Rodriguez, Yan Wang, Joseph Ozanne, Dogan Sumer, Dimitar Filev, and Damoon Soudbakhsh. Adaptive takeoff maneuver optimization of a sailing boat for america's cup. *Journal of Sailing Technology*, 7(01):88–103, 2022.

[137] Renato Rodriguez, Yan Wang, Joseph Ozanne, Jay Morrow, Dogan Sumer, Dimitar Filev, and Damoon Soudbakhsh. Adaptive learning and optimization of high-speed sailing maneuvers for america's cup. *Journal of Dynamic Systems, Measurement, and Control*, 145(2):021005, 2023.

[138] Dimitar Filev, Tomas Larsson, and Lixing Ma. Intelligent control for automotive manufacturing-rule based guided adaptation. In *IEEE IECON'00*, volume 1, pages 283–288, 2000.

[139] Jonathan H Tu. *Dynamic Mode Decomposition: Theory and Applications*. PhD thesis, Princeton University, 2013.

[140] Jennifer Annoni and Peter Seiler. A method to construct reduced-order parameter-varying models. *International Journal of Robust and Nonlinear Control*, 27(4):582–597, 2017.

[141] Kevin K Chen, Jonathan H Tu, and Clarence W Rowley. Variants of dynamic mode decomposition: boundary condition, koopman, and fourier analyses. *Journal of nonlinear science*, 22(6):887–915, 2012.

[142] Peter J Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of fluid mechanics*, 656:5–28, 2010.

[143] Clarence W Rowley, Igor Mezić, Shervin Bagheri, Philipp Schlatter, and Dan S Henningson. Spectral analysis of nonlinear flows. *Journal of fluid mechanics*, 641:115–127, 2009.

[144] Naoya Takeishi, Yoshinobu Kawahara, Yasuo Tabei, and Takehisa Yairi. Bayesian dynamic mode decomposition. In *IJCAI*, pages 2814–2821, 2017.

[145] Milan Korda and Igor Mezić. On convergence of extended dynamic mode decomposition to the koopman operator. *Journal of Nonlinear Science*, 28(2):687–710, 2018.

[146] Joshua L Proctor, Steven L Brunton, and J Nathan Kutz. Generalizing koopman theory to allow for inputs and control. *SIAM Journal on Applied Dynamical Systems*, 17(1):909–930, 2018.

[147] Steven L Brunton, Marko Budišić, Eurika Kaiser, and J Nathan Kutz. Modern koopman theory for dynamical systems. *arXiv preprint arXiv:2102.12086*, 2021.

[148] Maziar S Hemati, Clarence W Rowley, Eric A Deem, and Louis N Cattafesta. De-biasing the dynamic mode decomposition for applied koopman spectral analysis of noisy datasets. *Theoretical and Computational Fluid Dynamics*, 31(4):349–368, 2017.

[149] Igor Mezić. Analysis of fluid flows via spectral properties of the koopman operator. *Annual review of fluid mechanics*, 45:357–378, 2013.

[150] Dimitrios Giannakis. Data-driven spectral decomposition and forecasting of ergodic dynamical systems. *Applied and Computational Harmonic Analysis*, 47(2):338–396, 2019.

[151] Eurika Kaiser, J Nathan Kutz, and Steven L Brunton. Data-driven discovery of koopman eigenfunctions for control. *Machine Learning: Science and Technology*, 2(3):035023, 2021.

[152] Naoya Takeishi, Yoshinobu Kawahara, and Takehisa Yairi. Learning koopman invariant subspaces for dynamic mode decomposition. *Advances in Neural Information Processing Systems*, 30, 2017.

[153] Steven L Brunton, Bingni W Brunton, Joshua L Proctor, Eurika Kaiser, and J Nathan Kutz. Chaos as an intermittently forced linear system. *Nature communications*, 8(1):1–9, 2017.

[154] Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences*, 113(15):3932–3937, 2016.

[155] Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. Sparse identification of nonlinear dynamics with control (SINDYc). *IFAC-PapersOnLine*, 49(18):710–715, 2016.

[156] Urban Fasel, J Nathan Kutz, Bingni W Brunton, and Steven L Brunton. Ensemble-sindy: Robust sparse model discovery in the low-data, high-noise limit, with active learning and control. *Proceedings of the Royal Society A*, 478(2260):20210904, 2022.

[157] Eurika Kaiser, J Nathan Kutz, and Steven L Brunton. Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *Proceedings of the Royal Society A*, 474(2219):20180335, 2018.

[158] Giuseppe L'Erario, Luca Fiorio, Gabriele Nava, Fabio Bergonti, Hosameldin Awadalla Omer Mohamed, Emilio Benenati, Silvio Traversaro, and Daniele Pucci. Modeling, identification and control of model jet engines for jet powered robotics. *IEEE Robotics and Automation Letters*, 5(2):2070–2077, 2020.

[159] Jun Li, Gan Sun, Guoshuai Zhao, and H Lehman Li-wei. Robust low-rank discovery of data-driven partial differential equations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 767–774, 2020.

[160] Samuel Rudy, Alessandro Alla, Steven L Brunton, and J Nathan Kutz. Data-driven identification of parametric partial differential equations. *SIAM Journal on Applied Dynamical Systems*, 18(2):643–660, 2019.

[161] Chinmay S Kulkarni, Abhinav Gupta, and Pierre FJ Lermusiaux. Sparse regression and adaptive feature generation for the discovery of dynamical systems. In *International Conference on Dynamic Data Driven Application Systems*, pages 208–216. Springer, 2020.

[162] Li Yang and Abdallah Shami. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415:295–316, 2020.

[163] Renato Rodriguez, Omidreza Ahmadzadeh, Yan Wang, and Damoon Soudbakhsh. Physics-Informed and Temperature-Dependent Battery Digital Twin of State of Charge Dynamics. *Available at SSRN 4944568*, 2024.

[164] Shunli Wang, Carlos Fernandez, Yu Chunmei, Fan Yongcun, Cao Wen, Daniel-Ioan Stroe, and Zonghai Chen. *Battery System Modeling*. Elsevier, 2021.

[165] Caiping Zhang, Jiuchun Jiang, Yang Gao, Weige Zhang, Qiujiang Liu, and Xiaosong Hu. Charging optimization in lithium-ion batteries based on temperature rise and charge time. *Applied energy*, 194:569–577, 2017.

[166] D. Anseán, M. Dubarry, A. Devie, B. Y. Liaw, V. M. García, J. C. Viera, and M. González. Fast charging technique for high power LiFePO4 batteries: A mechanistic analysis of aging. *J. Power Sources*, 321:201–209, 2016.

[167] Peter Keil and Andreas Jossen. Charging protocols for lithium-ion batteries and their impact on cycle life—An experimental study with different 18650 high-power cells. *Journal of Energy Storage*, 6:125–141, 2016.

[168] Ruoyun Shi, Sepehr Semsar, and Peter W. Lehn. Constant current fast charging of electric vehicles via a DC grid using a dual-inverter drive. *IEEE Transactions on Industrial Electronics*, 64(9):6940–6949, 2017.

[169] Muhammad Usman Tahir, Ariya Sangwongwanich, Daniel-Ioan Stroe, and Frede Blaabjerg. Overview of multi-stage charging strategies for Li-ion batteries. *J. Energy Chem.*, 2023.

[170] Christian Campestrini, Stephan Kosch, and Andreas Jossen. Influence of change in open circuit voltage on the state of charge estimation with an extended Kalman filter. *J. Energy Storage*, 12:149–156, 2017.

[171] Xiaosong Hu, Dongpu Cao, and Bo Egardt. Condition monitoring in advanced battery management systems: Moving horizon estimation using a reduced electrochemical model. *IEEE/ASME Trans. Mechatronics*, 23(1):167–178, 2017.

[172] Huazhen Fang, Yebin Wang, and Jian Chen. Health-aware and user-involved battery charging management for electric vehicles. *IEEE Trans. Control Syst. Technol.*, 25(3):911–923, 2016.

[173] Saehong Park, Donggun Lee, Hyoung Jun Ahn, Claire Tomlin, and Scott Moura. Optimal control of battery fast charging based-on Pontryagin's minimum principle. In *IEEE Conference CDC'2020*, pages 3506–3513. IEEE, 2020.

[174] Goran Kujundžić, Šandor Ileš, Jadranko Matuško, and Mario Vašak. Optimal charging of valve-regulated lead-acid batteries based on model predictive control. *Applied Energy*, 187:189–202, 2017.

[175] Suryanarayana Kolluri, Sai Varun Aduru, Manan Pathak, Richard D. Braatz, and Venkat R. Subramanian. Real-time nonlinear model predictive control (NMPC) strategies using physics-based models for advanced lithium-ion battery management system (BMS). *Journal of The Electrochemical Society*, 167(6):063505, 2020.

[176] Ji Liu, Guang Li, and Hosam K Fathy. An extended differential flatness approach for the health-conscious nonlinear model predictive control of lithium-ion batteries. *IEEE Trans. Control Syst. Technol.*, 25(5):1882–1889, 2016.

[177] Reinhardt Klein, Nalin A Chaturvedi, Jake Christensen, Jasim Ahmed, Rolf Findeisen, and Aleksandar Kojic. Optimal charging strategies in lithium-ion battery. In *ACC'2021*, pages 382–387. IEEE, 2011.

[178] Renato Rodriguez, Yan Wang, Joseph Ozanne, Dogan Sumer, Dimitar Filev, and Damoon Soudbakhsh. Adaptive takeoff maneuver optimization of a sailing boat for america's cup. *J. Sail. Technol.*, 7(01):88–103, 2022.

[179] Dimitar P Filev, Sunil Bharitkar, and Meng-Fu Tsai. Nonlinear control of static systems with unsupervised learning of the initial conditions. In *18th NAFIPS'1999 (Cat. No. 99TH8397)*, pages 169–173, 1999.

[180] Renato Rodriguez, Yan Wang, Joseph Ozanne, Jay Morrow, Dogan Sumer, Dimitar Filev, and Damoon Soudbakhsh. Adaptive learning and optimization of high-speed sailing maneuvers for america's cup. *J. Dyn. Sys., Meas., Control*, 145(2):021005, 2023.

[181] Gianluca Manduca, Zhaoxuan Zhu, Polina B. Ringler, Guodong Fan, and Marcello Canova. Model Order Reduction of the Doyle-Fuller-Newman Model via Proper Orthogonal Decomposition and Optimal Collocation. In *2023 IEEE Vehicle Power and Propulsion Conference (VPPC)*, pages 1–6. IEEE, 2023.

[182] Yanpeng Li, Yiwei Feng, Chuang Wang, Ziwen Xing, Dawei Ren, and Lin Fu. Performance evaluation and optimization of the cascade refrigeration system based on the digital twin model. *Applied Thermal Engineering*, page 123160, 2024.

[183] James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2), 2012.

[184] Samuel Rudy, Alessandro Alla, Steven L Brunton, and J Nathan Kutz. Data-driven identification of parametric partial differential equations. *SIAM J. Appl. Dyn.*, 18(2):643–660, 2019.

[185] Aoxue Chen and Guang Lin. Robust data-driven discovery of partial differential equations with time-dependent coefficients. *arXiv preprint arXiv:2102.01432*, 2021.

[186] Daniel A. Messenger and David M. Bortz. Weak SINDy For Partial Differential Equations. *Journal of Computational Physics*, 443:110525, October 2021.

[187] N Shawki, R Rodriguez Nunez, I Obeid, and J Picone. On automating hyperparameter optimization for deep learning applications. In *2021 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pages 1–7. IEEE, 2021.

[188] Matthew Mackay, Paul Vicol, Jonathan Lorraine, David Duvenaud, and Roger Grosse. Self-Tuning Networks: Bilevel Optimization of Hyperparameters using Structured Best-Response Functions. In *International Conference on Learning Representations*, 2018-09.

[189] Daniel Jakubovitz, Raja Giryes, and Miguel RD Rodrigues. Generalization error in deep learning. In *Compressed Sensing and Its Applications: Third International MATHEON Conference 2017*, pages 153–193. Springer, 2019.

[190] Renato Rodriguez, Yan Wang, Jozeph Ozanne, Dogan Sumer, Dimitar Filev, and Damoon Soudbakhsh. Adaptive learning for maximum takeoff efficiency of high-speed sailboats. *IFAC-PapersOnLine*, 55(12):402–407, 2022.

[191] Kiarash Movassagh, Arif Raihan, Balakumar Balasingam, and Krishna Pattipati. A critical look at coulomb counting approach for state of charge estimation in batteries. *Energies*, 14(14):4074, 2021.

[192] Kong Soon Ng, Chin-Sien Moo, Yi-Ping Chen, and Yao-Ching Hsieh. Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries. *Applied energy*, 86(9):1506–1511, 2009.

[193] M. Derakhshan, E. Sahraei, and D. Soudbakhsh. Detecting mechanical indentation from the time constants of li-ion batteries. *Cell Reports Physical Science*, 2022 (In press).

[194] Ping-Chun Tsai, Bohua Wen, Mark Wolfman, Min-Ju Choe, Menghsuan Sam Pan, Liang Su, Katsuyo Thornton, Jordi Cabana, and Yet-Ming Chiang. Single-particle measurements of electrochemical kinetics in NMC and NCA cathodes for Li-ion batteries. *Energy & Environmental Science*, 11(4):860–871, 2018.

[195] Ruifeng Zhang, Bizhong Xia, Baohua Li, Libo Cao, Yongzhi Lai, Weiwei Zheng, Huawen Wang, Wei Wang, and Mingwang Wang. A study on the open circuit voltage and state of charge characterization of high capacity lithium-ion battery under different temperature. *Energies*, 11(9):2408, 9 2018.

[196] BK Purushothaman, PW Morrison, and U Landau. Reducing mass-transport limitations by application of special pulsed current modes. *J. Electrochem. Soc.*, 152(4):J33, 2005.