Physics Informed Gated Recurrent Unit (GRU) Network for SOC & SOH Prediction in Lithium-Ion Batteries

Vishnu Ram Jawaharram
1233423648

Yashaswini Inala
1220247028

Zarin Musarrat Manita
1234504871

I. Introduction & Background

The modern world runs on battery power. The world's most critical industries are widely adopting lithium-ion (Li-ion) battery-powered systems, but battery performance depends on how well they're managed [1]. State of Charge (SOC) and State of Health (SOH) estimation algorithms for Battery Management Systems (BMS) don't just monitor battery safety, they provide the actionable insights needed to manage risk and extend a system's operational life [2]. The SOC and SOH of a Li-ion battery cannot be measured; it has to be estimated from the measurable parameters such as the battery voltage, current, and temperature.SOC indicates the remaining energy left to use in a Li-ion battery at that particular instant of time.SOH is defined as the remaining useful life of a battery when compared to when it was actually new. More precise SOC and SOH data empower engineers with better control over battery performance, reliability, and cost-effectiveness. Nonetheless, achieving consistent and accurate battery state estimation remains difficult due to the variety of battery chemistry and the complex, varying operating environments they face. This challenge requires advanced methods that consider diverse battery conditions to enable reliable monitoring and management

Usually, SOC and SOH are estimated using either the Coulomb Counting (CC) method or the Capacity Fade method. The CC-based SOC estimation, shown in Eq. 1, accumulates error over time due to the constant term involved in the integral. There are several drawbacks associated with the CC method. To improve CC accuracy, the initial battery SOC must be known from the Open Circuit Voltage (OCV) vs. SOC curve provided by the manufacturer. However, the OCV of a battery is not directly measurable. To obtain the OCV vs. SOC curve, the battery is tested by charging and discharging at a very low C-rate[3].

$$SOC(t) = SOC(t_0) - \frac{1}{Q_{nom}} \int_{t_0}^{t} I(\tau)\, d\tau \tag{1}$$

- Q_{nom}: The *nominal capacity* of the battery, typically measured in Ampere-hours (Ah). It represents the maximum charge the battery can store as specified by the manufacturer.
- $SOC(t_0)$: The *initial state of charge* at the starting time t_0. It reflects the fraction or percentage of battery capacity available at the beginning of the measurement.

The actual battery is often modeled using a Pseudo-Two-Dimensional (P2D) electrochemical model, which involves numerous parameters and requires solving complex partial differential equations (PDEs). This high computational demand makes it impractical for implementation in onboard Battery Management Systems (BMS) [4]. Alternatively, batteries are modeled using Electrical Equivalent Circuit (EQC) models. These can be first-order, second-order, or third-order models, depending on the accuracy required. EQC models are commonly used onboard within the actual BMS software components (SWC) due to their lower complexity and computational efficiency.

The data-driven approach has become a highly discussed research topic in battery management algorithms in recent years. Although it can achieve a high order of accuracy, the resulting SOC or SOH predictions are purely data-driven, treating the neural network as a black box without explicit physical insight into the system. This paper aims to integrate a physics-based loss function into the neural network, introducing physical constraints directly into the learning process. The physics-based loss is derived from either the P2D model or the equivalent circuit model, providing the network with knowledge of the underlying physics governing SOC estimation and thereby improving its reliability and [5].

II. Method

LSTMs and GRUs, while effective at modeling time-series data, do not incorporate any physical laws or equations that directly influence the estimation of battery state. Their predictions are purely data-driven and lack interpretability regarding the underlying battery physics. There are a lot of works involving LSTMs and GRUs for battery State of charge estimation, clearly the GRUs have a better performance in terms of the accuracy when compared to the former[6].

Physics-Informed Neural Networks (PINNs), on the other hand, are specifically designed to embed physical laws—often represented as partial differential equations (PDEs)—directly into the learning process. In addition to the standard data-driven loss function used in multilayer perceptrons (MLPs), PINNs include a physics-based loss term that enforces the network's outputs to satisfy known physical constraints. The core architecture of a PINN is typically a fully connected feedforward neural network (MLP), which is trained to minimize both the data loss and the physics-based loss. This approach enables the network to learn solutions that are not only accurate with respect to the data but also consistent with the governing physical principles of the system. Previous work has involved combining LSTM and PINNS to estimate battery degradation modeling and SOH [7]. The flow in which the project is going to proceed is shown in Fig.1.

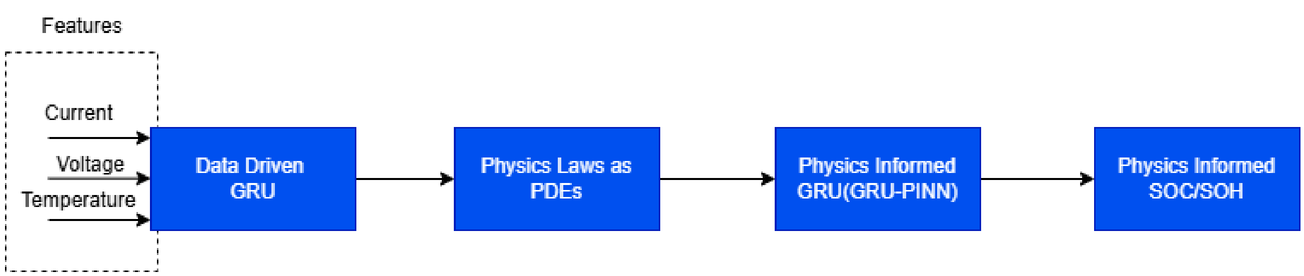

Figure 1: *Block diagram of the proposed battery SOC/SOH estimation approach.*

In this project, a GRU neural network will first be developed as a baseline model for battery SOC and SOH estimation, using a purely data-driven approach based on the dataset described in the next section. Next, a physics-based loss function will be incorporated to guide SOC estimation according to the OCV relationship EQC or P2D models. For SOH estimation, the model will also include PDEs that describe battery degradation mechanisms, such as Loss of Lithium Inventory (LLI), Loss of Active Anode Material (LAMA), and Loss of Active Cathode Material (LAMC). These degradation processes can be characterized using Incremental Capacity Analysis (ICA) and Differential Voltage Analysis (DVA), providing additional physics-based constraints to improve the interpretability and reliability of the network's predictions [8].

Expected Advantage of Combining GRU with PINNs

GRUs are highly effective in modeling time-series data, enabling the network to learn complex dependencies of the system's state based on its previous states with respect to BMS-measurable parameters such as current, voltage, and temperature. This capability is essential for capturing the dynamic behavior of batteries. As discussed, GRUs often outperform LSTMs for battery state estimation tasks, offering greater efficiency and comparable or better accuracy. Extending this approach by integrating a physics-based loss function during training—thus forming a PINN—ensures that the model's predictions are not only data-driven but also consistent with established battery physics, thereby improving reliability and interpretability. This should also give a greater edge in scaling this model to different parameters and also might be useful in predicting the SOC of Lithium Ferrous Phosphate(LFP) battery chemistry since the OCV vs SOC curve is flat across different voltage ranges, making it hard to predict with all those uncertainty in actual hardware systems accuracy.

III. Datasets and Resources

The following datasets are decided on for the project:

- **XJTU dataset** [9] – This dataset contains 55 commercial 18650 Li-ion cells cycled under six different charge/discharge protocols, with voltage, current, and capacity data.
- **MIT dataset** [10] – Dataset of 124 Li-ion cells cycled under fast-charging protocols; includes voltage, current, capacity, and cycle life data. Used in the Nature Energy paper of the same title.
- **TJU dataset** [11] – This contains records voltage relaxation profiles from commercial 18650 cells, enabling capacity estimation through relaxation behavior.
- **HUST dataset** [12] – This dataset provides over 140,000 cycles, enabling research on personalized SoH/RUL prediction under varied load conditions.

The following software and tools can be utilized:

- **General Python Libraries:**
  - `Pandas`: Data manipulation, cleaning, and preprocessing.
  - `NumPy`: Numerical operations.
  - `Scikit-learn`: Feature extraction, scaling, normalization, and splitting data into training and testing sets.
- **Deep Learning Frameworks:**
  - `PyTorch`: A widely used deep learning framework for model training.
  - `TensorFlow`: Another deep learning framework for building and training models.
  - `Simulink`: A simulation Platform to validate the model, given that we get the same cell's data.
- **Libraries and Tools:**
  - `Scikit-learn`: For clustering algorithms like K-means, DBSCAN, and hierarchical clustering.
  - `UMAP` (Uniform Manifold Approximation and Projection): For dimensionality reduction to visualize high-dimensional data.
  - `t-SNE` (t-Distributed Stochastic Neighbor Embedding): A technique for reducing dimensions and visualizing clustering results.
  - `Matplotlib`: A plotting library for 2D visualizations.
  - `Plotly`: For interactive visualizations of clustering and feature analysis.
- **Tools for Tracking and Collaboration:**
  - `TensorBoard`: TensorFlow's tool for visualizing model performance and experiment tracking.
  - `GitHub`: For version control, collaboration, and sharing of code.
  - `Jupyter Notebooks`: An interactive environment for Python code.
- **Tools for Documentation and Reporting:**
  - `LaTeX`: For preparing a detailed technical report or academic paper.

Brief Timeline

- **10/1 - 10/15**: Literature review and dataset pre-processing
- **10/16 - 10/31:** Learning model development and initial experiments
- **10/28:** Midterm deliverables with preliminary results and insights from literature review
- **11/1 - 11/15:** Model refinement and performance evaluations
- **11/16 - 11/30:** Rough draft documentation and final presentation preparation
- **12/1 onward (Finals week):** Final report completion and submission

Tasks and Responsibilities

This is a broad division of tasks:

- **Vishnu Ram Jawaharram:** Will work on Data Preprocessing and initial experiments.
- **Yashaswini Inala:** Will focus on the development of the learning model
- **Zarin Musarrat Manita:** Will work on literature review and performance evaluation.

References

[1] Gavin Harper, Roberto Sommerville, Emma Kendrick, Laura Driscoll, Peter Slater, Rustam Stolkin, Allan Walton, Paul Christensen, Oliver Heidrich, Simon Lambert, et al. Recycling lithium-ion batteries from electric vehicles. *nature*, 575(7781):75–86, 2019.

[2] Volodymyr Andrushchak. Why accurate soc and soh estimation is critical for battery longevity and business success. *Tech Briefs*, August 2025.

[3] Yunhong Che, Le Xu, Remus Teodorescu, Xiaosong Hu, and Simona Onori. Enhanced soc estimation for lfp batteries: A synergistic approach using coulomb counting reset, machine learning, and relaxation. *ACS Energy Letters*, 10(2):741–749, 2025.

[4] Pouya Hashemzadeh, Martin Désilets, Marcel Lacroix, and Ali Jokar. Investigation of the p2d and of the modified single-particle models for predicting the nonlinear behavior of li-ion batteries. *Journal of Energy Storage*, 52:104909, 2022.

[5] Minzhi Chen, Guijun Ma, Weibo Liu, Nianyin Zeng, and Xin Luo. An overview of data-driven battery health estimation technology for battery management system. *Neurocomputing*, 532:152–169, 2023.

[6] Mouleshwar Saravanan, J Ajaykumar, K S Arun, S P Sibhiecharan, and T Ananthan. Lithium-ion battery soc prediction using deep learning. In *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–6, 2024.

[7] Yan Ding, Jinqi Zhu, Yang Liu, Dan Ning, and Mingyue Qin. Integrated framework of lstm and physical-informed neural network for lithium-ion battery degradation modeling and prediction. *AI*, 6(7), 2025.

[8] Xiaojun Li, David Jauernig, Mengzhu Gao, and Trevor Jones. 7 - battery cloud with advanced algorithms. In Mohammadreza Daneshvar, Behnam Mohammadi-Ivatloo, Kazem Zare, and Amjad Anvari-Moghaddam, editors, *IoT Enabled Multi-Energy Systems*, pages 111–136. Academic Press, 2023.

[9] Fujin Wang, Zhenyu Zhai, Zhenyu Zhao, et al. Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis. *Nature Communications*, 15:4332, 2024.

[10] Keith A. Severson, Peter M. Attia, Norman Jin, Nicholas Perkins, Bryan Jiang, Zhiwei Yang, Michael H. Chen, Muratahan Aykol, Peter K. Herring, Dimitrios Fraggedakis, Martin Z. Bazant, Stephen J. Harris, William C. Chueh, and Richard D. Braatz. Data-driven prediction of battery cycle life before capacity degradation, 2019.

[11] Jiangong Zhu. Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation, 2022.

[12] Ye Yuan, Guijun Ma, and Songpei Xu. Hust battery dataset: 77 lfp/graphite cells under 77 personalized discharge protocols, 2022.