

Predictive parametric inference of lithium-ion battery degradation trajectory through high-throughput physics-based simulation

Peiheng Lai

The Hong Kong Polytechnic University

Chenhang Zheng

The Hong Kong Polytechnic University

Hao Wu

The Hong Kong Polytechnic University

Hao Cheng

The Hong Kong Polytechnic University

Bing Li

Harbin Institute of Technology

Viktor Korzhikov-Vlakh

Saint-Petersburg State University

Xiaohe Song

Shenzhen Eacomp Technology Co., Ltd

Minggang Zeng

zengmg@a-star.edu.sg

Institute of Advanced Intelligence and Computing

Ming Yang

The Hong Kong Polytechnic University

Article

Keywords:

Posted Date: September 24th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11005119/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Inferring the physical parameters that govern lithium-ion battery degradation is critical for the prediction of batter lifetime, which however remains challenging owing to the complexity of physics-based models, the non-uniqueness of inverse solutions, and the experimental inaccessibility of governing parameters. Here, we develop a physics-based simulation emulator framework for predictive parametric inference of battery degradation trajectories. A neural-network emulator, trained on a dataset generated by pseudo-two-dimensional (P2D) battery simulations, allows rapid forward evaluation from parameters to degradation trajectories and thereby replaces costly repeated simulation. This high-throughput capability enables Bayesian inversion that maps experimental trajectories onto the most probable latent parameters among the candidate solutions, while incorporating physical constraints and similarities among cycling conditions. Most significantly, the inferred parametric representation provides a physics-grounded basis for forecasting complete degradation trajectories. Using only the early-life 100 – 95% state-of-health segment, our approach reduces the RMSE of late-life trajectory prediction to 59.00 cycles, in contrast with 79.62 cycles by the best data-driven approach reported. Together, these results establish an efficient route from physics-based parametric inference to practical degradation prediction, particularly where experimental data are scarce.

Introduction

Lithium-ion batteries (LIBs) are a cornerstone of modern energy storage systems, especially in electric vehicles (EVs) and renewable-energy applications^{1–3}. Their widespread use makes reliable lifetime prediction, with aging experimental data, increasingly important for the safety evaluation and energy management of LIBs. However, battery aging experiments are inherently time-consuming and resource-intensive, which require long cycling periods, strictly controlled conditions, and repeated measurements across numerous cells^{4,5}. These demands make it costly to assess battery performance across broad design and operational spaces. As a result, accurate and efficient lifetime prediction is essential for accelerating battery development and reducing experimental burden.

Efforts to evaluate and interpret LIBs degradation mainly follow two routes: data-driven modeling with deep learning (DL) and physics-based simulation (PBS). Data-driven deep learning can achieve high lifetime-prediction accuracy when large and representative dataset are available^{6,7}. Such models nevertheless depend heavily on training data and transfer poorly across operational conditions. In contrast, physics-based simulation offers a mechanistic route to degradation. Pseudo-two-dimensional (P2D) modelling coupled with ageing sub-models describes electrochemical transport, reaction kinetics and side reactions in explicit physical terms^{8,9}. This makes physics-based simulation well suited for understanding degradation from its underlying processes rather than through macroscopic behaviour alone. Most degradation-related parameters, however, cannot be measured or validated directly in experiments. Moreover, multiple coupled mechanisms can produce similar degradation behaviour,

making it difficult to distinguish and quantify their individual contributions^{10–12}. These issues complicate comprehensive correspondence between simulation outcomes and experimental observations.

In recent years, there has been increasing interest in physics-guided deep learning. By incorporating physical principle and prior knowledge into machine learning frameworks, such approaches can improve performance while preserving mechanistic consistency in battery lifetime prediction. Physics-informed neural networks (PINNs) are a representative example, where governing equations, typically a system of partial differential equations (PDEs), are incorporated into the training objective to constrain the model^{13,14}. However, PINNs face several challenges that limit their broader application. Training instability and nonconvergence remain persistent problems. PDE-constrained training is highly sensitive to the specification of loss function, initial and boundary conditions, which are often difficult to define precisely in real experimental systems¹⁵. These challenges motivate practical approaches that maintain the physical fidelity of degradation modeling while leveraging the efficiency of deep learning.

In this study, we propose a physics-based simulation emulator framework for battery degradation analysis and prediction. The framework uses a deep learning emulator to represent the parameter-to-trajectory mapping generated by PBS. Bayesian inference then matches experimental degradation trajectories and estimates latent degradation parameters in the latent physical space. Early-to-late degradation prediction further predicts later-life behavior from latent representations inferred from early-life observation. These integrated components connect experimental observations to degradation mechanisms while reducing the need for computational-costly simulation and optimization.

Results

Physics-based simulation emulator

We first constructed a physics-based simulation emulator as a surrogate model to accelerate the repeated forward evaluations required in optimization and inference (see the framework in Fig. 1), as directly solving the full PBS model for every candidate parameter set would be costly in computational resources and time. The emulator is a neural network (NN) that learns the nonlinear mapping from degradation parameters to corresponding aging trajectories, trained on data generated by PBS (Fig. 1b). Because it reproduces the PBS solution across the defined parameter space, the emulator enables high-throughput evaluation and screening of large parameter spaces without repeated simulation.

The PBS dataset was generated by sampling six degradation-governing parameters within predefined ranges (Table 1). These parameters covered major degradation pathways (Fig. 1a), including solid electrolyte interphase (SEI) growth, lithium plating and dead-lithium decay, and loss of active material (LAM)^{16,17}. The parameters and the associated electrochemical equations are described in **Supplementary Note 1**.

Table 1

Key degradation parameters and sampling ranges for the physics-based simulation dataset generation. Six governing parameters were varied around their baseline values using Latin hypercube sampling (LHS) to generate the training dataset for the PBS emulator. Variation ranges are given as multipliers applied to the baseline value. For example, a range of [5.0, 30] denotes sampled values between 5 and 30 times the baseline.

Parameter	Symbol	Baseline value	Variation
Initial SEI thickness [m]	L_{SEI}	5.0×10^{-9}	[1.0, 2.0]
SEI solvent diffusivity [$\text{m}^2 \text{s}^{-1}$]	D_{sol}	2.5×10^{-22}	[5.0, 30]
Lithium plating kinetic rate constant [ms^{-1}]	k_{pl}	1.0×10^{-9}	[10, 50]
Dead lithium decay constant [s^{-1}]	γ	1.0×10^{-6}	[10, 100]
NE critical stress [Pa]	σ_{cr}	6.0×10^7	[0.80, 1.20]
NE LAM constant proportional term [s^{-1}]	β_{stress}	4.0×10^{-6}	[1.0, 3.0]

In total, 400 successful physics-based simulations of degradation from 100 to 80% state-of-health (SOH) were generated. As shown in Fig. 2a, the PBS dataset covers a broad range of degradation phenotypes, from fast-fading to long-lived trajectories. The sampled PBS parameter space produced distinct degradation modes that are different in lifetimes and curvature patterns. The dataset was partitioned (8:2) into training and testing set for emulator development and evaluation. On the test set (Fig. 2b), the emulator accurately reproduced the PBS-generated trajectories by predicting cycle numbers at integer SOH levels with RMSE of 0.0118 in normalized cycle space (12.95 cycles on the original scale) and R^2 of 0.9942. Prediction errors increase in the long lifetime region, which is sparsely represented in the PBS dataset. These results demonstrate that the emulator can capture and reproduce the nonlinear parameter-to-trajectory mapping with accuracy and efficiency needed for high-throughput screening and inference.

High-throughput Bayesian full-trajectory inference

We next developed a high-throughput Bayesian inference workflow to map degradation trajectories in the TONGJI public battery dataset¹⁸ (Fig. 3a) into the PBS-emulated parameter space. The objective was to identify candidate parameter sets whose emulated degradation trajectories best matched the shape of experimental trajectory. For each experimental target, the emulator allows hundreds of PBS-derived candidate trajectories to be evaluated across the parameter space. However, similar trajectory shapes

can arise from different parameter combinations, leaving the inversed inference solution non-unique. Hence, we applied Bayesian inference to impose additional structure on the candidates and to identify the most probable parameter set within the PBS-emulated latent physical space (Fig. 1c). Experimental and PBS-emulated trajectories were normalized by their own cycle ranges before shape mapping, so that the inference emphasized relative degradation shape and pacing. The inferred parameters are therefore treated as latent and effective degradation descriptors for relative comparison across cells and conditions, rather than as experimentally observable physical quantities with absolute values.

The Bayesian model introduced multiple sources of structure into the candidate screening¹⁹, as illustrated in Fig. 1e. For each experimental trajectory S_i , a candidate pool was first drawn from the high-throughput emulator search. Each parameter candidate θ_{ik} was assigned a posterior-like probability ω_{ik} by:

$$\omega_{ik} \propto p(S_i | \theta_{ik}) \cdot p_{\text{group}}(\theta_{ik}) \cdot p_{\text{dens}}(\theta_{ik}) \cdot p_T(\theta_{ik} | T_i) \# (1)$$

The trajectory-fitting likelihood $p(S_i | \theta_{ik})$ rewards candidates through trajectory matching. The group prior $p(\theta_{ik})$ promotes clustering of cells with similar operating conditions in the parameter space. The density prior $p_{\text{dens}}(\theta_{ik})$ penalizes candidates with largest nearest-neighbor distance in the parameter space. The temperature term $p_T(\theta_{ik} | T_i)$ restricted parameters with Arrhenius temperature dependence^{20,21}. For instance, SEI growth is expected to be stronger and lithium plating weaker at higher temperature. The candidate with the highest probability was selected as the inferred latent parameter set for the specific cell. Details of Bayesian inference are provided in **Supplementary Note 2**.

The high-throughput fitting strategy accurately mapped experimental degradation trajectories to PBS-emulated candidates (Fig. 3). Fitted and experimental cycle positions agreed closely across all SOH levels (Fig. 3b), with an RMSE of 0.0104 in normalized cycle space (RMSE = 13.03 cycles on the original scale) and $R^2 = 0.9957$.

We further examined the inferred latent parameters to assess the physical consistency and interpretability of the inference. For the temperature-dependent degradation trajectories in TONGJI dataset 2, the inferred parameters show clear and physically consistent trends. Electrochemical understanding and experimental observations both indicate that increasing temperature promotes SEI growth and SEI-related side reactions^{22,23}, consistent with the increases in initial SEI thickness and SEI solvent diffusivity that we infer (Fig. 4a-b). This mainly contributes to early-stage capacity loss. Higher temperature can also accelerate the striping of plated lithium and reduce the tendency for lithium plating^{24,25}. Meanwhile, the inferred LAM-related parameters vary less with temperature, showing an increase in negative electrode stress and a decrease in the LAM proportional term (Fig. 4e-f). Because LAM governs mainly late-stage degradation, its weaker contribution at higher temperature is consistent with the slower late-stage fading and longer cycle life observed in the high-temperature group^{26,27}. For TONGJI dataset 3, the discharge rate was varied from 1C to 4C at a fixed temperature with minor changes in degradation trajectory and curvature, and the inferred parameters therefore exhibited much

smaller differences between protocols. Together, these results show that the physics-constrained Bayesian inference selects effective parameter sets consistent with both experimental trajectory shapes and expected degradation physics.

Early-to-late degradation prediction

Finally, we assessed the practical applicability of the PBS-emulated inference for degradation prediction in a data-scarce context, where only limited aging trajectories are available. Here, the emulator serves as a physically grounded mapping between latent parameters and full degradation trajectories. As demonstrated in Fig. 1d, early-stage observations were first mapped into the PBS-emulated space to infer most probable latent parameters for each cell. The inferred parameters were then passed through the emulator to reconstruct the complete degradation trajectory, enabling late-life prediction from early-life data.

The prediction followed a group-anchor manifold strategy. For each operating-condition group, one complete experimental cell was used as an anchor to define and constrain the condition-specific candidate searched during early-life inference of testing cells (**Supplementary Note 3**). The remaining 28 cells were used as test cells, and only their early-life trajectory segments were provided for inference. We evaluated three independent prediction windows: 1) applying the observed 100 – 95% SOH segment to predict the subsequent 94 – 80% SOH trajectory, 2) using 100 – 90% SOH segment to predict 89 – 80% SOH segment, and 3) using 100 – 85% SOH segment to predict 84 – 80% SOH segment (Fig. 5a-c). We repeated the entire prediction procedure across five independent split trials and reported the averaged performance metrics.

We compared our method with conventional data-driven approaches, including CNN, LSTM, and Transformer models. For full-trajectory prediction (Fig. 5d), the PBS emulator performed comparably to the best data-driven model in the shorter 85–80% and 90–80% windows, with RMSEs (in cycles) of 33.71 (against the best data-driven result 26.52 by LSTM) and 53.96 (against the best data-driven result 45.69 by CNN). Its advantage becomes clearer in the most difficult 95 – 80% prediction window, where PBS achieves the lowest RMSE of 59.00 cycles, outperforming RMSE of 79.62 cycles from the best data-driven model (Transformer). Endpoint cycle prediction at 80% SOH shows a similar trend (Fig. 5e). The PBS emulator is close to the best model in the 85 – 80% task, with a MAPE of 5.2% against 4.5% for LSTM. In the 90 – 80% and 95 – 80% prediction windows, the PBS emulator produces the lowest MAPE of 7.8% and 11.7%, (compared with the best data-driven results of 8.6% and 13.9%), respectively. PBS also shows more stable prediction performance (smaller standard deviation in prediction error, represented by shorter error bar) in these prediction tasks. These results indicate that the framework is well suited to degradation prediction in data-limited settings, where physical priors can compensate for a lack of training data.

Discussion

In this study, we have developed a PBS-emulator framework for battery degradation analysis, effective parameter inference, and early-to-late prediction. Physics-based simulations were first applied to generate a degradation space constrained by electrochemical-aging mechanisms. A neural-network emulator learned the mapping from degradation-related parameters to SOH-cycle trajectories, enabling rapid forward evaluation without repeated P2D simulations. The emulator further supported Bayesian full-trajectory inference of latent parameters and early-to-late prediction from limited experimental data. The framework shows how physical priors can compensate for a scarcity of complete ageing trajectories.

Compared with direct PINN-based inverse modeling, the PBS-emulator strategy provides a more effective and stable computational route for battery degradation inference. PINN approaches require repeated optimization under coupled PDE constraints and can suffer from unstable convergence and difficult loss balancing. In contrast, the main computational cost of the present framework is shifted to the generation of PBS trajectories. In our implementation, approximately 4 successful simulations (about 1670 cyclic aging cycles) were generated per hour on a single CPU core. Once the PBS dataset is generated, subsequent tasks run at speeds comparable to conventional forward data-driven models. This separation between offline physics-based data generation and fast online inference makes the framework attractive for high-throughput screening, parameter-space exploration, and degradation prediction in data-limited battery studies.

A key limitation of this framework is that fitting, inference, and prediction all depend on the coverage and fidelity of the generated PBS space. The emulator can only interpolate within the degradation behaviors represented by the latent parameter space, and experimental trends outside this space cannot be fully captured. In our results, the PBS-derived trajectories deviated from the accelerated fading observed in the 25°C subgroup of TONGJI dataset 2, indicating a limitation of the current simulation rather than only an inference error. This discrepancy may reflect missing or under-represented degradation pathways, imperfect parameter ranges, or protocol-specific effects not fully captured by the PBS model. Future work should refine the degradation model and test the framework on more experimental datasets and battery chemistries to establish its general applicability.

Methods

Pseudo-two-dimensional simulation

The pseudo-two-dimensional (P2D)⁸ simulations for battery degradation in this study were implemented by open-source software PyBaMM²⁸. The aging simulation was implemented using a cyclic schedule comprising three sequential steps: 1) constant-current charging at 1C rate to 4.2V, 2) constant-voltage hold at 4.2V until a cut-off current at 1/20C, 3) constant current discharging at 1C to 2.5V. The ambient temperature is set to 25°C.

Evaluation Metric

RMSE, MAPE and coefficient of determination R^2 are chosen to evaluate the prediction performance of DL models. Given a group of N samples with observed cycle life y_i and predicted cycle life $\hat{y}_i$, RMSE (in cycles) is defined as

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \# (2)$$

MAPE (in %) is defined as

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|\hat{y}_i - y_i|}{y_i} \times 100\% \# (3)$$

The coefficient of determination R^2 is defined as

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \# (4)$$

Declarations

Competing interests

The authors declare no competing interests.

Author Contribution

P. H. L.: Writing – original draft, Methodology, Investigation, Formal analysis, Validation, Data curation, Visualization. C. H. Z.: Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation. H. W., H. C., B. L., V. K.V., and X. H. S.: Writing – review & editing, Methodology, Investigation. M. G. Z.: Writing – review & editing, Validation, Formal analysis, Supervision. M. Y.: Writing – review & editing, Conceptualization, Supervision, Project administration, Funding acquisition. All authors modified and approved the manuscript.

Acknowledgements

M. Y. acknowledges the funding support from Guangdong Natural Science Foundation (project number: 2024A1515010031) and Hong Kong Polytechnic University (project number: P0042711, P0058378, P0058408, P0058414, P0053568, and P0048122). M.G.Z. acknowledges funding support from the A*STAR MTC programmatic funding (Grant No.: M23L9b0052) and the National Supercomputing Center (NSCC) Singapore for the use of the high-performance computing facilities.

Data Availability

The experimental battery dataset in this study is from TONGJI dataset 2 and dataset 318, which can be directed downloaded from <https://doi.org/10.5281/zenodo.6379165>. A BatteryML29 preprocessed data version can be obtained from BatteryLife30 in <https://zenodo.org/records/19688272>. All code implementations including data processing, model training and result visualization are available at <https://github.com/LPHQaq/Predictive-parametric-inference-of-LIBs>.

Code Availability

All code implementations including data processing, model training and result visualization is available at <https://github.com/LPHQaq/Predictive-parametric-inference-of-LIBs>.

References

1. Faruque, O., Isalm, M. J., Faruq, O. & Alam, M. M. A comprehensive review of past and present developments of Li-ion batteries. *Discov. Sustain.* **7**, 881 (2026).
2. Ngoy, K. R. *et al.* Lithium-ion batteries and the future of sustainable energy: A comprehensive review. *Renew. Sustain. Energy Rev.* **223**, 115971 (2025).
3. Zhang, B. *et al.* Lithium-ion battery recycling relieves the threat to material scarcity amid China's electric vehicle ambitions. *Nat. Commun.* **16**, 6661 (2025).
4. Sun, Y. *et al.* Designing safe and long-life lithium-ion batteries via a solvent-relay strategy. *Nat. Energy* **10**, 1450–1457 (2025).
5. Luh, M. & Blank, T. Comprehensive battery aging dataset: capacity and impedance fade measurements of a lithium-ion NMC/C-SiO cell. *Sci. Data* **11**, 1004 (2024).
6. Zhang, H. *et al.* Battery lifetime prediction across diverse ageing conditions with inter-cell deep learning. *Nat. Mach. Intell.* **7**, 270–277 (2025).
7. Hu, J. *et al.* Early prediction of lithium-ion battery degradation with a generative pre-trained transformer. *Nat. Commun.* **17**, 126 (2025).
8. Doyle, M., Fuller, T. F. & Newman, J. Modeling of Galvanostatic Charge and Discharge of the Lithium/Polymer/Insertion Cell. *J. Electrochem. Soc.* **140**, 1526 (1993).
9. Jokar, A., Rajabloo, B., Désilets, M. & Lacroix, M. Review of simplified Pseudo-two-Dimensional models of lithium-ion batteries. *J. Power Sources* **327**, 44–55 (2016).
10. Khalik, Z., Donkers, M. C. F., Sturm, J. & Bergveld, H. J. Parameter estimation of the Doyle–Fuller–Newman model for Lithium-ion batteries by parameter normalization, grouping, and sensitivity analysis. *J. Power Sources* **499**, 229901 (2021).
11. Xu, L., Lin, X., Xie, Y. & Hu, X. Enabling high-fidelity electrochemical P2D modeling of lithium-ion batteries via fast and non-destructive parameter identification. *Energy Storage Mater.* **45**, 952–968 (2022).

12. Zheng, C. *et al.* Inverse inference of lithium-ion battery degradation parameters via physical modeling and data-driven machine learning. *Npj Comput. Mater.* <https://doi.org/10.1038/s41524-026-02250-3> (2026) doi:10.1038/s41524-026-02250-3.
13. Karniadakis, G. E. *et al.* Physics-informed machine learning. *Nat. Rev. Phys.* **3**, 422–440 (2021).
14. Luo, K. *et al.* Physics-informed neural networks for PDE problems: a comprehensive review. *Artif. Intell. Rev.* **58**, 323 (2025).
15. Wang, S., Teng, Y. & Perdikaris, P. Understanding and Mitigating Gradient Flow Pathologies in Physics-Informed Neural Networks. *SIAM J. Sci. Comput.* **43**, A3055–A3081 (2021).
16. Li, R. *et al.* The importance of degradation mode analysis in parameterising lifetime prediction models of lithium-ion battery degradation. *Nat. Commun.* **16**, 2776 (2025).
17. Edge, J. S. *et al.* Lithium ion battery degradation: what you need to know. *Phys. Chem. Chem. Phys.* **23**, 8200–8221 (2021).
18. Zhu, J. *et al.* Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation. *Nat. Commun.* **13**, 2261 (2022).
19. van de Schoot, R. *et al.* Bayesian statistics and modelling. *Nat. Rev. Methods Primer* **1**, 1 (2021).
20. Wang, K. *et al.* Temperature-dependent electrochemical process decoupling in PEM water electrolysis: DRT deconvolution and Arrhenius-based modeling. *Appl. Energy* **407**, 127351 (2026).
21. Chung, Y. & Green, W. H. New modified Arrhenius equation to describe the temperature dependence of liquid phase reaction rates. *Chem. Eng. J.* **516**, 163300 (2025).
22. Wu, Y., Ge, G., Wang, S., Xiong, L. & He, Z. Formation mechanisms of solid electrolyte interphase and its influence on lithium battery performance. *Mater. Today Energy* **54**, 102124 (2025).
23. Rodrigues, M.-T. F., Sayed, F. N., Gullapalli, H. & Ajayan, P. M. High-temperature solid electrolyte interphases (SEI) in graphite electrodes. *J. Power Sources* **381**, 107–115 (2018).
24. Xu, G. *et al.* Visualized Detection of Lithium Plating on Graphite Anodes Cycled Under Low Temperature and Fast Charging. *Adv. Funct. Mater.* **35**, 2412614 (2025).
25. Konz, Z. M. *et al.* High-throughput Li plating quantification for fast-charging battery design. *Nat. Energy* **8**, 450–461 (2023).
26. Peng, Y. *et al.* Quantitative Analysis of Active Lithium Loss and Degradation Mechanism in Temperature Accelerated Aging Process of Lithium-Ion Batteries. *Adv. Funct. Mater.* **34**, 2404495 (2024).
27. Ebner, M., Marone, F., Stampanoni, M. & Wood, V. Visualization and Quantification of Electrochemical and Mechanical Degradation in Li Ion Batteries. *Science* **342**, 716–720 (2013).
28. Sulzer, V., Marquis, S. G., Timms, R., Robinson, M. & Chapman, S. J. Python Battery Mathematical Modelling (PyBaMM). *J. Open Res. Softw.* **9**, 14–14 (2021).
29. Zhang, H. *et al.* BatteryML: An Open-source Platform for Machine Learning on Battery Degradation. in *International Conference on Learning Representations* vol. 2024 14808–14836 (2024).

30. Tan, R. et al. BatteryLife: A Comprehensive Dataset and Benchmark for Battery Life Prediction. in *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2* 5789–5800 (Association for Computing Machinery, New York, NY, USA, 2025).
doi:10.1145/3711896.3737372.

Figures

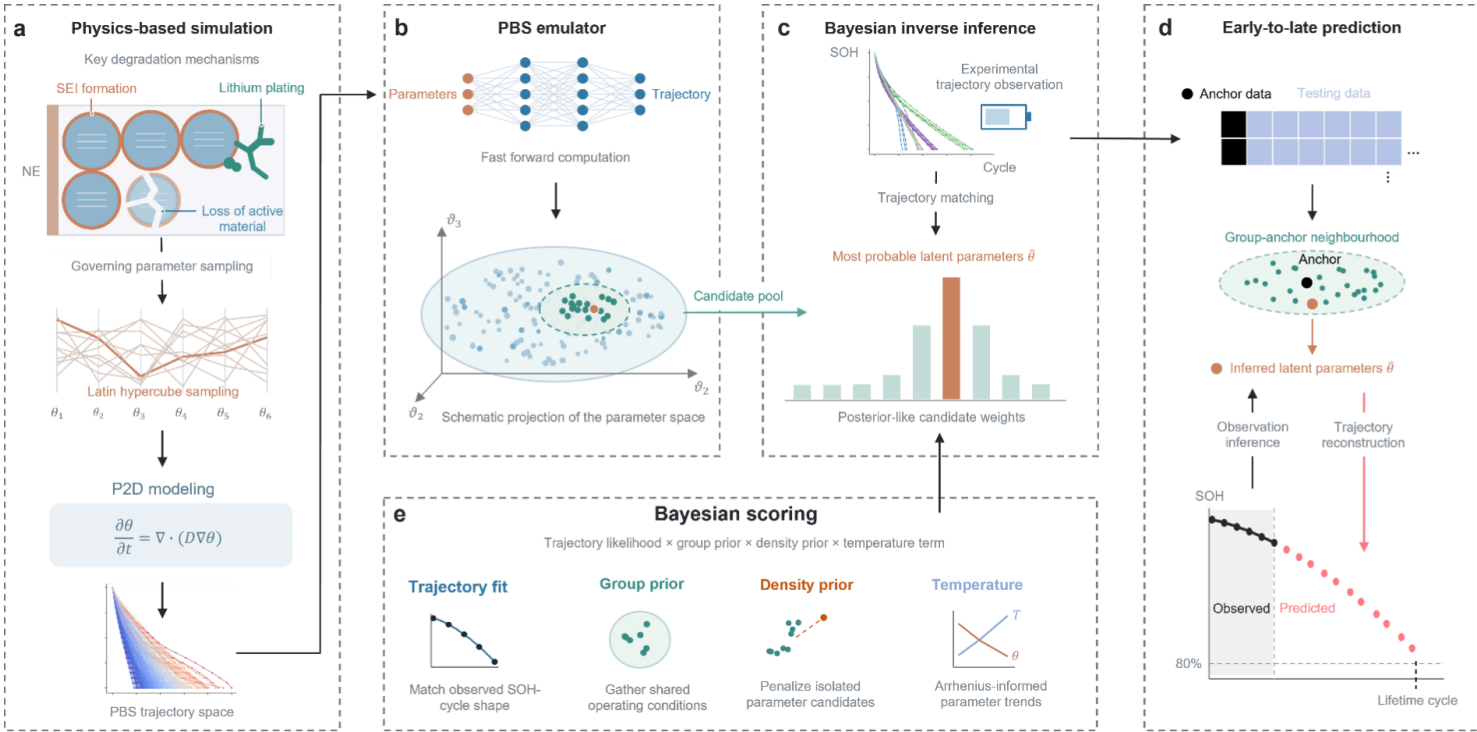


Figure 1

Overview of predictive parametric inference for battery degradation trajectories. **a.** Physics-based P2D simulations are employed to generate degradation trajectories, wherein the range of key degradation parameters are specified in Table 1. **b.** The simulated trajectories are used to train a forward PBS emulator. **c.** Experimental trajectories are mapped to the emulator space through Bayesian inverse inference to obtain the most probable degradation parameters. **d.** Early-life observations are combined with anchor information to predict full degradation trajectories from the inferred parameters. **e.** Bayesian scoring regularizes trajectory matching using shape likelihood, grouping, location density, and temperature-dependent priors.

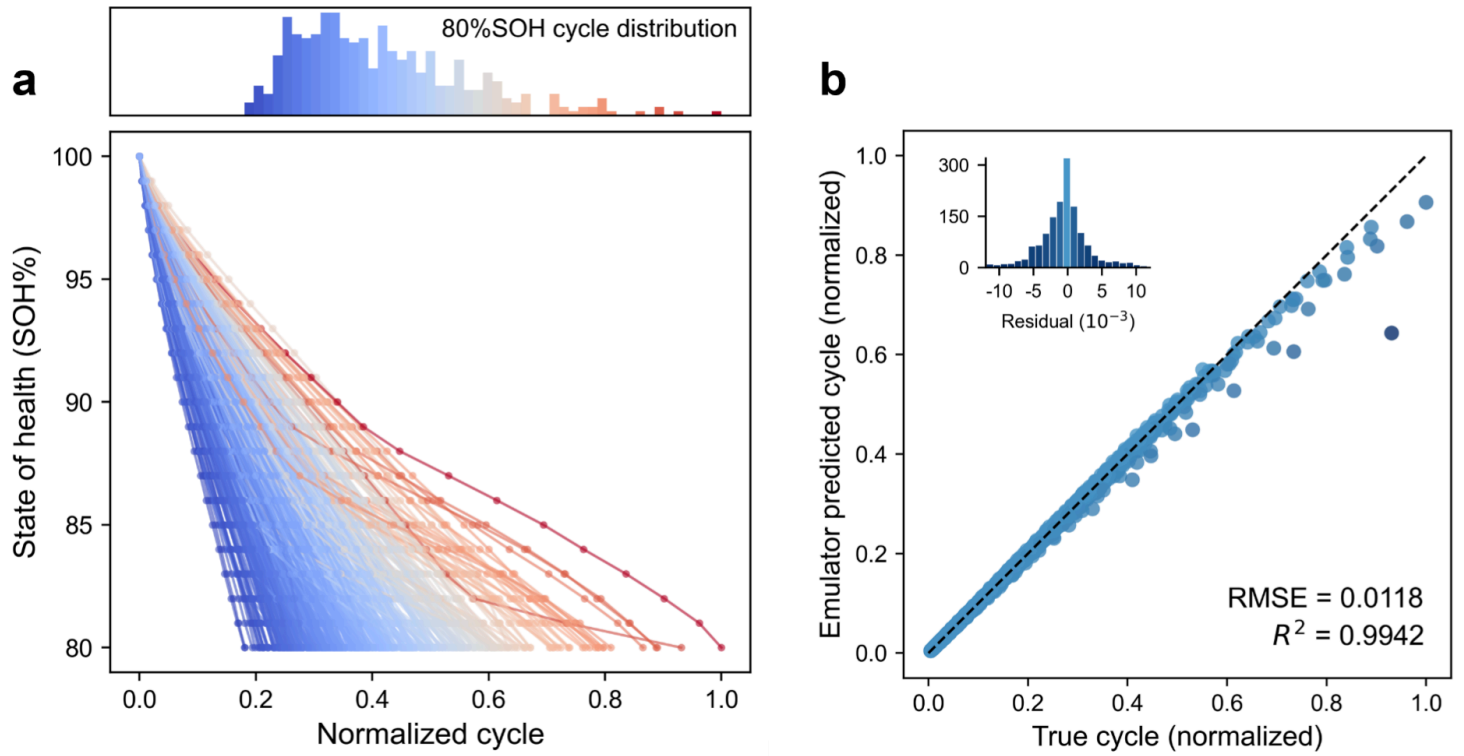


Figure 2

Physics-based simulation emulator. a. PBS-generated degradation trajectories represented by state-of-health (SOH) versus cycle number. Curves are plotted in normalized cycle space (normalized by 1089 cycle) and colored by lifetime at 80% SOH. The upper histogram shows the lifetime distribution. **b.** Emulator performance on the PBS test set, comparing predicted and true normalized cycles at integer SOH points. The dashed line indicates perfect agreement. RMSE and R^2 are calculated for evaluation.

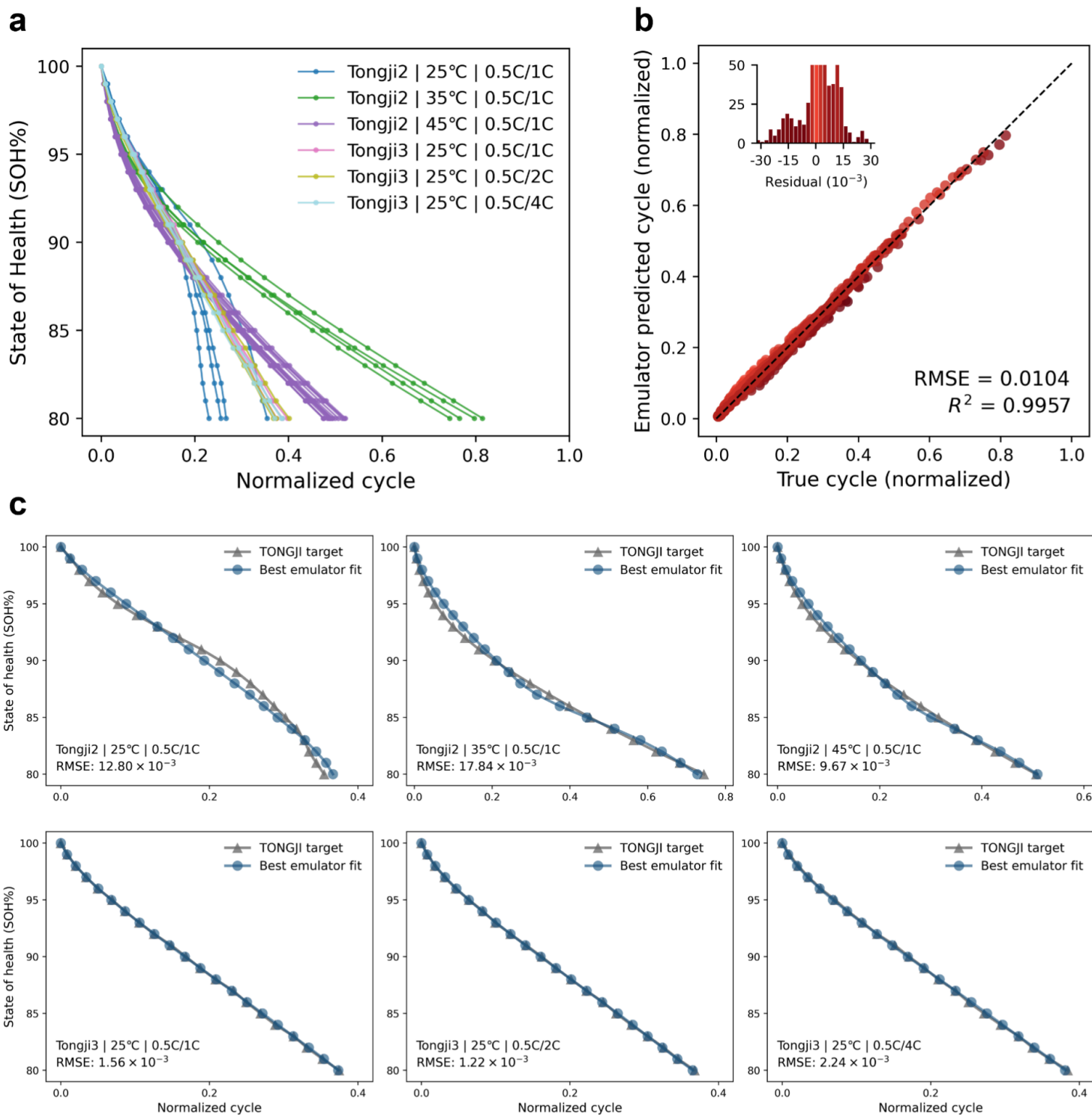


Figure 3

High-throughput full-trajectory fitting performance. **a.** Normalized SOH-cycle trajectories from the Tongji public battery dataset. Cells were cycled under different temperatures and charge/discharge-rate protocols, including TONGJI dataset 2 with cells at 25, 35, and 45 °C under 0.5/1C cycling and TONGJI dataset 3 with cells at 25 °C under 0.5/1C, 0.5/2C, and 0.5/4C cycling (i.e. charge/discharge rate). Dataset 2 comprises nickel manganese cobalt (NMC) batteries, and dataset 3 comprises batteries blend

of nickel cobalt aluminum (NCA) and NMC material. **b.** Agreement between experimental target trajectories and fitted PBS-emulated trajectories in normalized cycle space. The dashed line indicates perfect agreement, and the inset figure shows the residual distribution. **c.** Representative best full-trajectory fits for each operating condition. Gray trajectories denote experimental targets and blue trajectories the best PBS emulator fits.

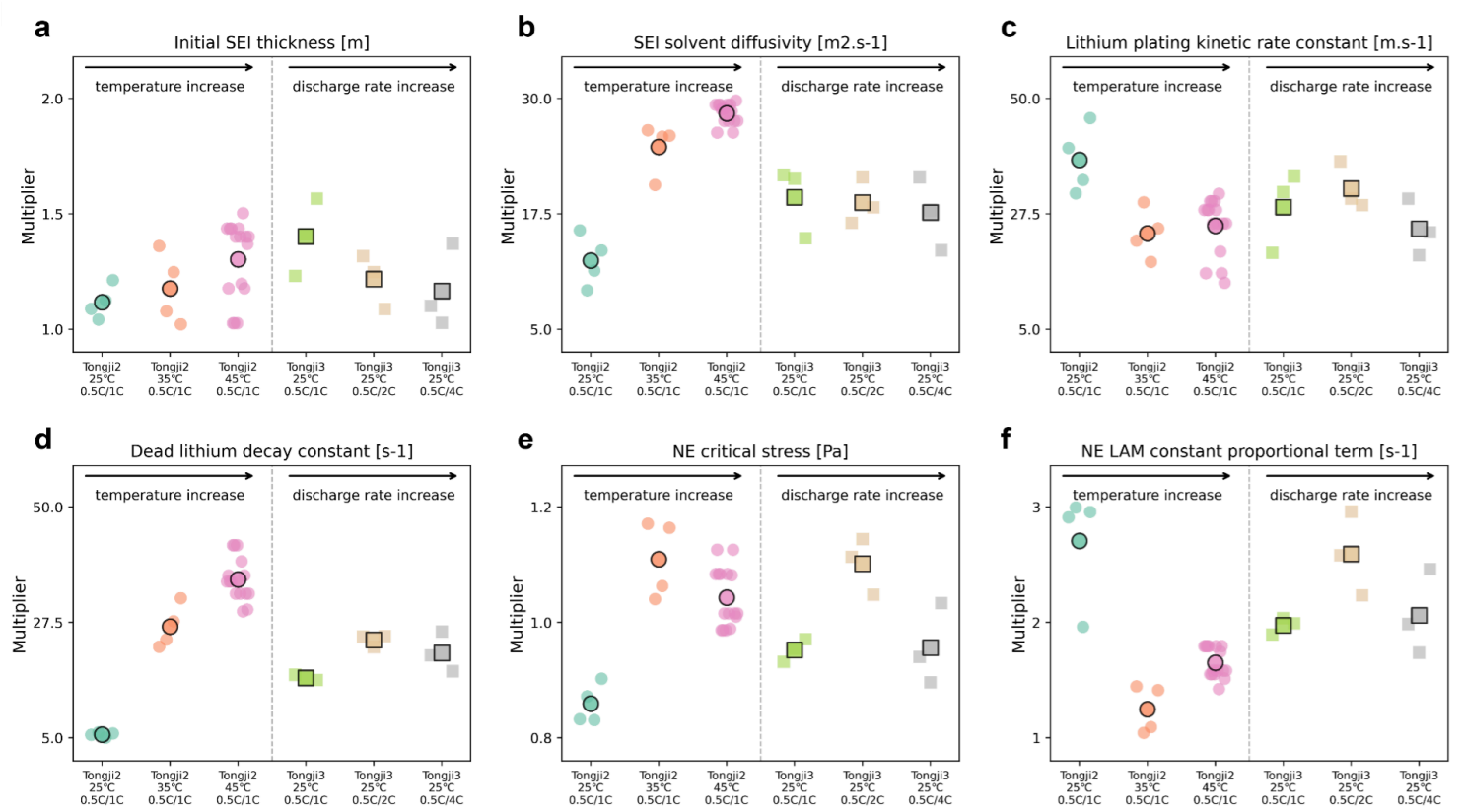


Figure 4

Bayesian inference of latent degradation parameters. Inferred parameter multipliers (compared to baseline value in **Table 1**) for experimental trajectories under different operating conditions. Each small marker represents inferred result of one experimental cell, and larger outlined markers indicate group-level means. Panels show latent multipliers for **a-b.** SEI growth, **c-d.** lithium plating and dead-lithium decay, and **e-f.** LAM.

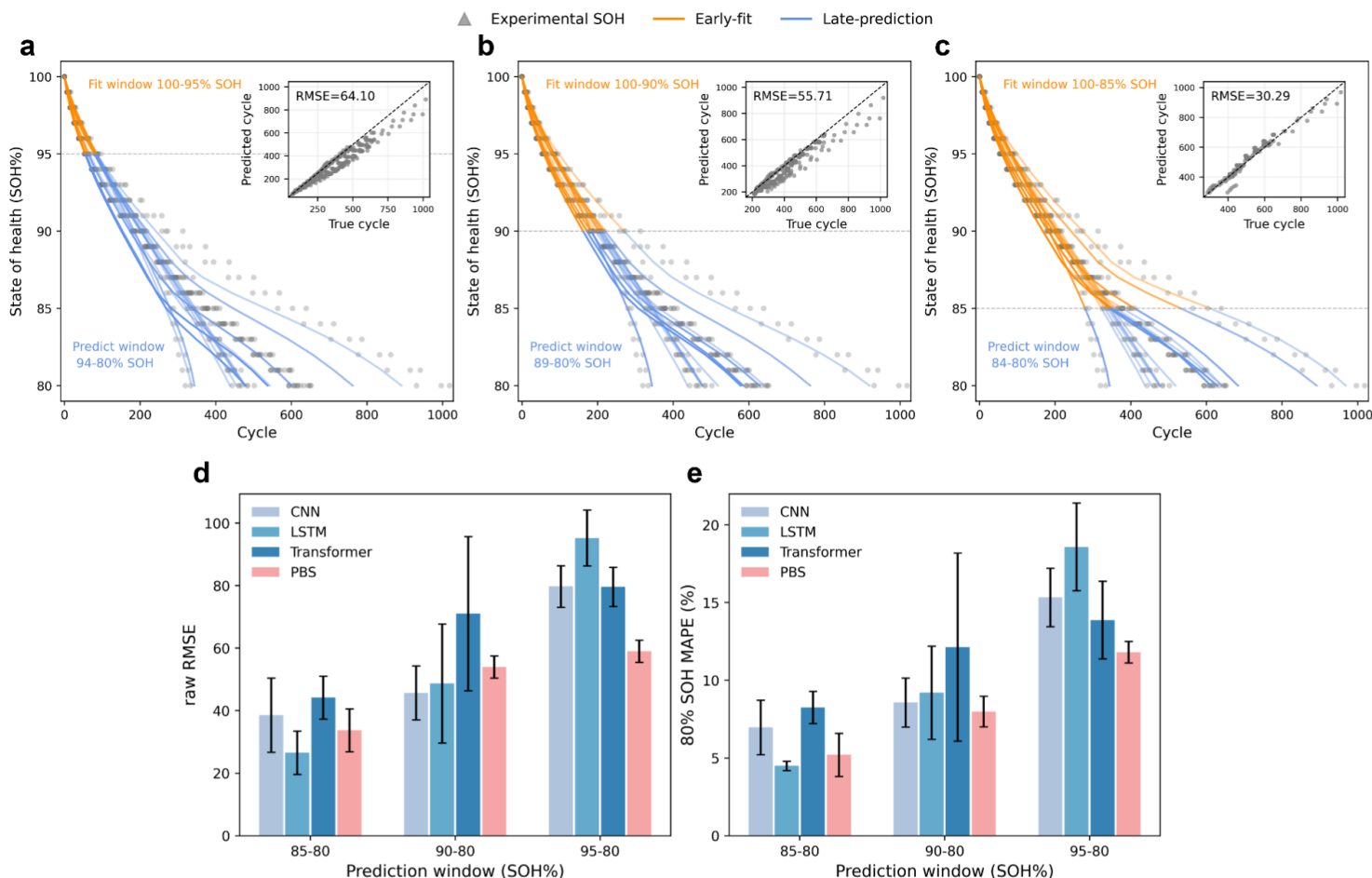


Figure 5

Early-to-late degradation prediction with PBS-emulated inference. a-c. PBS-emulator predictions for three early-life input windows. Orange curves indicate observed early-life fitting segment, blue curves the predicted late-life trajectory and grey points the experimental target. Insets compare predicted and true cycle positions on the original scale. **d.** Trajectory-level RMSE (original scale) and **e.** MAPE for the predicted cycle at 80% SOH, comparing the PBS emulator with the data-drive approaches.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SM.docx](#)