






# Remaining Useful Life Prediction of Lithium-Ion Battery With Adaptive Noise Estimation and Capacity Regeneration Detection

Jiushi Zhang , Student Member, IEEE, Yuchen Jiang , Member, IEEE, Xiang Li , Student Member, IEEE, Hao Luo , Senior Member, IEEE, Shen Yin , Senior Member, IEEE, and Okayay Kaynak, Life Fellow, IEEE

## I. INTRODUCTION

### A. Research Background and Related Work

**Abstract**—As an indispensable energy device, 18650 lithium-ion battery has widespread applications in electric vehicles. Remaining useful life (RUL) prediction of lithium-ion battery is critical for the normal operation of electric vehicles. In conventional approaches, the adaptive estimation of model parameters and the detection of capacity regeneration await further research. To adaptively estimate the noise variables in the degradation model and to accurately detect the battery capacity regeneration, this article proposes a novel expectation maximization-unscented particle filter-Wilcoxon rank sum test (EM-UPF-W) approach. In detail, in the case of unlabeled small samples, this article constructs a dynamic degradation model on the basis of UPF for a single battery, which adaptively estimates the noise variables with the aid of EM algorithm. Furthermore, the Wilcoxon rank sum test is introduced to determine the capacity regeneration point, so as to decrease the prediction error. A 18650 lithium-ion battery dataset produced by NASA is used to demonstrate the approach. Experimental results show that the proposed EM-UPF-W outperforms some existing data-driven techniques.

**Index Terms**—Adaptive noise estimation, battery, capacity regeneration, expectation maximization-unscented particle filter (EM-UPF), prediction, remaining useful life (RUL).

With the popularization of green transportation systems, as an intelligent traffic infrastructure, new energy electric vehicles have developed rapidly in recent years, to solve the problems such as energy substitution and environmental pollution [1], [2], [3], [4]. The 18650 lithium-ion battery has become a vital component in the battery system of electric vehicles with the advantages of long cycle life, high energy density, and wide operational temperature range attributable to the long cycle life, high energy density, and wide operational temperature range, etc [5], [6], [7], [8]. With the accumulation of usage time, the capacity of battery gradually decreases, which reduces the reliability of electric vehicles. Consequently, the accurate prediction of the time elapsed from the current moment to the failure of the battery, remaining useful life (RUL), is essential for the safety of the electric vehicles [9], [10], [11].

In general, the lithium-ion battery RUL prediction approaches include operational degradation mechanism-based methods, data-driven methods, and hybrid empirical methods [12]. Considering the degradation process of lithium-ion battery is complicated, the operational degradation mechanism-based methods have notable limitations [13], [14]. On the other hand, the data-driven approaches do not demand prior knowledge of the battery's operational mechanism, the RUL can be predicted from the data generated during the working cycles. Nevertheless, these learning-based algorithms are black-box models, which demand massive historical data to train algorithm models [15], [16], [17], [18]. Furthermore, the uncertainty of the battery during degradation cannot be quantified [12].

Hybrid empirical-based RUL prediction approaches have strong interpretability and reliability. By incorporating prior knowledge into the approach, the algorithm model between the battery health index and the parameters, such as current, voltage, and capacity is constructed. Moreover, Kalman filter (KF) [12], and particle filter (PF) [19], are used for parameter identification, so that the battery's RUL is predicted. Bhaskar et al. proposed a Bayesian framework through equivalent circuit method for battery's RUL prediction, which is a pioneering practical prognosis approach [20]. Considering the influence of

Manuscript received 12 April 2022; revised 15 June 2022 and 28 July 2022; accepted 25 August 2022. Date of publication 9 September 2022; date of current version 18 April 2023. Recommended by Technical Editor Zhiwei Gao and Senior Editor Zhiwei Gao. This work was supported in part by the National Natural Science Foundation of China under Grant 62073104 and in part by the China Postdoctoral Science Foundation under Grant 2022M710965. (Corresponding author: Hao Luo.)

Jiushi Zhang, Yuchen Jiang, Xiang Li, and Hao Luo are with the Department of Control Science and Engineering, Harbin Institute of Technology, Harbin 150001, China (e-mail: hit\_zjs@163.com; yc.jiang@hit.edu.cn; lixianghit@yeah.net; hao.luo@hit.edu.cn).

Shen Yin is with the Department of Mechanical and Industrial Engineering, Faculty of Engineering, Norwegian University of Science and Technology, 7034 Trondheim, Norway (e-mail: shen.yin@ntnu.no).

Okayay Kaynak is with the Department of Electrical and Electronics Engineering, Bogazici University, 34342 Istanbul, Turkey (e-mail: kaynak@boun.edu.tr).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TMECH.2022.3202642>.

Digital Object Identifier 10.1109/TMECH.2022.3202642

circulating current, Liu et al. constructed a capacity degradation model. In addition, PF algorithm is introduced to identify the parameters of the model to predict the battery's RUL [21]. Toward the integration of the prior knowledge in RUL prediction, Guha et al. proposed a lithium-ion battery capacity degradation model that combines PF algorithm and battery internal resistance growth model, which has a great RUL prediction performance compared with the conventional single model [22]. To improve the traditional PF algorithm, Miao et al. introduced unscented PF into the RUL prediction of the battery, so as to acquire excellent prediction performance [19]. Li et al. incorporated a data-driven least squares support vector machine into the UPF model, which adopts an online iterative strategy to promote the RUL prediction accuracy. Nevertheless, when most of the existing UPF-based approaches are used in battery RUL prediction, the noise variables of the state transition equation and measurement equation require human setting. There is a lack of adaptive noise variables with the addition of capacity data, which requires in-depth research [23].

During the operational process, the battery's capacity gradually decreases with the accumulation of usage in general. Nevertheless, the electrochemical performance may resume shortly when the battery is in a static state, so that the battery capacity is regenerated [24]. In this sense, suddenly increased battery capacity will have a large error in the RUL prediction. Therefore, accurate determination of the capacity regeneration point is of great significance. Researchers have done some work on the detection and treatment of capacity regeneration. Benjamin et al. proposed a particle-filtering-based for battery's RUL prediction, which can consider the influence of capacity regeneration phenomena [25]. In [26] and [27], when predicting the RUL, the capacity regeneration points are removed, which may lose crucial degradation information. Considering capacity regeneration, Park et al. proposed an long short-term memory (LSTM)-based battery RUL prediction with multiple channel charging profiles [28]. Ma et al. proposed a particle filter-autoregressive-based (PF-AR) RUL prediction approach with capacity regeneration detection, the elimination of false positives in detection process needs further research [29]. To separate the local regeneration and the global degradation during the battery capacity degradation, Pang et al. proposed a multiscale wavelet decomposition method [30]. In addition, these two parts are combined by constructing an artificial neural network for RUL prediction. To comprehensively describe the battery degradation process, Zhao et al. divided the battery's degradation into capacity regeneration and normal degradation. The predictions are modified by alternating grey model and relevance vector machine to improve accuracy [31], [32].

### B. Motivations and Contributions of This Work

It should be noted that data-driven RUL prediction methods demand the online data follow a similar distribution law to the offline data, which require massive historical data to train the algorithm model [12]. Conversely, in practice, especially for the newly installed 18650 lithium-ion batteries in electric vehicles, the acquisition of historical data for numerous batteries of the

same category has notable limitations [17], [33]. Furthermore, the research on capacity regeneration needs in-depth work.

Based on these observations, the motivations of this article can be outlined as follows:

- 1) The noise variables of the state transition equation and measurement equation based on hybrid empirical approaches demand human setting. How to realize adaptive noise estimation requires in-depth research [12], [31].
- 2) How to accurately detect the capacity regeneration point of the battery, and give the statistical boundary between the capacity regeneration points and the noncapacity regeneration points, thereby decreasing the RUL prediction error needs to be further studied [26], [27], [28], [29], [30].

In this sense, this article proposes a novel expectation maximization-unscented particle filter-Wilcoxon rank sum test (EM-UPF-W) approach to predict the RUL of the battery. The contributions of this article can be summarized in following these folds.

- 1) Without a large number of historical data, a dynamic UPF model is proposed to describe the degradation of a single battery, from which the uncertainty of the capacity and the RUL during degradation can be represented. Furthermore, the noise variables in the UPF model can be well adaptively estimated based on EM algorithm.
- 2) An approach based on Wilcoxon rank sum test is proposed to measure the difference between the prior probability density function (PDF) and posterior PDF of the particles in UPF model, from which the capacity regeneration points can be accurately detected, so as to decrease the error in RUL prediction.
- 3) A 18650 lithium-ion battery dataset is used to demonstrate the effectiveness of the proposed EM-UPF-W approach. Experimental results show that the proposed approach outperforms some existing techniques.

The rest of the article is organized as follows. The next section formulates the problem. Section III introduces the methodology of RUL prediction approach. Section IV analyses a 18650 lithium-ion battery dataset. Finally, Section V concludes this article.

## II. PROBLEM FORMULATION

Compared with the single exponential model and the linear model, the double exponential model can better fit the capacity degradation process of lithium-ion battery [34]. Consequently, the double exponential model is introduced to describe the capacity degradation process, shown as follows:

$$q_k = a_k e^{b_k \cdot k} + c_k e^{d_k \cdot k} \quad (1)$$

where  $k$  is the index of the battery operational cycles,  $a_k$ ,  $b_k$ ,  $c_k$ , and  $d_k$  are the model parameters that need to be identified,  $q_k$  is the battery capacity under the  $k$ th operational cycle. *The research problem in this paper is to construct a battery degradation model for RUL prediction on the basis of the EM-UPF only using the battery capacity data, in which the capacity regeneration points are detected.*

It should be noted that root mean square error (RMSE) and mean absolute error (MAE) are used to evaluate the RUL prediction performance, which can be written in (2)–(3)

$$\text{RMSE} = \sqrt{\frac{1}{K} \sum_{k=1}^K (\text{RUL}_{pk} - \text{RUL}_{tk})^2} \quad (2)$$

$$\text{MAE} = \frac{1}{K} \sum_{k=1}^K |\text{RUL}_{pk} - \text{RUL}_{tk}| \quad (3)$$

where  $K$  is the total number of operational cycles,  $\text{RUL}_{pk}$  is the predicted RUL value, and  $\text{RUL}_{tk}$  is the RUL ground truth of the  $k$ th cycle, respectively. The small RMSE and MAE indicate the excellent RUL prediction performance.

### III. METHODOLOGY OF RUL PREDICTION APPROACH

#### A. Adaptive State Update Model Based on UPF Algorithm

Compared with conventional KF and EKF, PF can handle the parameter identification of nonlinear system [34]. It should be noted that the prior PDF is used as the importance function in PF algorithm. Nevertheless, the particles generated from the prior PDF may not reflect the true particle distribution. Furthermore, the current measured information of the system is ignored. Based on these observations, UPF is introduced for adaptive state update model. Specifically, the posterior PDF of the unscented Kalman filter (UKF) is used as the prior PDF of PF. The battery degradation model based on UPF can be written as follows:

$$\begin{cases} x_k = x_{k-1} + u_k \\ y_k = h(x_k) + v_k \end{cases} \quad (4)$$

where  $x_k = [a_k, b_k, c_k, d_k]^T$ ,  $u_k = [u_a, u_b, u_c, u_d]^T$ ,  $v_k \in R^{1 \times 1}$ . The state transition in (4) can be written in (5)

$$\begin{cases} a_k = a_{k-1} + u_a & (u_a \sim N(0, \sigma_a^2)) \\ b_k = b_{k-1} + u_b & (u_b \sim N(0, \sigma_b^2)) \\ c_k = c_{k-1} + u_c & (u_c \sim N(0, \sigma_c^2)) \\ d_k = d_{k-1} + u_d & (u_d \sim N(0, \sigma_d^2)) \end{cases} \quad (5)$$

where  $u_a, u_b, u_c, u_d$  are the noise terms for state variables, respectively,  $\sigma_a^2, \sigma_b^2, \sigma_c^2, \sigma_d^2$  are the variances of the corresponding noise terms, respectively. Correspondingly, the measurement equation can be shown in (6)

$$y_k = a_k e^{b_k \cdot k} + c_k e^{d_k \cdot k} + v_k \quad (v_k \sim N(0, \sigma_v^2)) \quad (6)$$

where  $v_k$  is the measured noise, the variance of which is  $\sigma_v^2$ . The necessary steps of the UPF algorithm are described as follows:

##### 1) Particle initialization:

According to the initial state variable  $x_0$ , the particle set,  $\{r_0^{(i)} | i = 1, 2, \dots, N_s\}$ , is generated with the initial distribution of the state variables  $p(x_0)$ , where  $N_s$  is the number of particles. The initial weight of each particle is  $\frac{1}{N_s}$ . The initial mean  $\bar{r}_0$  and the initial variance  $S_0$  of these particles are written in (7) and (8)

$$\bar{r}_0 = E(r_0) \quad (7)$$

$$S_0 = E[(r_0 - \bar{r}_0)(r_0 - \bar{r}_0)^T] \quad (8)$$

##### 2) State update based on UKF:

After the  $k$ -1th iteration, the Sigma points of each particle can be constructed, as shown follows:

$$\chi_{k-1}^{(i)} = [\bar{r}_{k-1}^{(i)}, \bar{r}_{k-1}^{(i)} + \sqrt{\alpha S_{k-1}^{(i)}}, \bar{r}_{k-1}^{(i)} - \sqrt{\alpha S_{k-1}^{(i)}}] \quad (9)$$

where the number of the Sigma points of each particle is  $2n+1$ ,  $i$  is the index of these particles, the corresponding weight of which can be shown in (10)–(12), respectively:

$$W_{0,m}^{(i)} = \frac{\lambda}{n + \lambda} \quad (10)$$

$$W_{0,c}^{(i)} = \frac{\lambda}{n + \lambda} (1 + \beta - \alpha^2) \quad (11)$$

$$W_{j,m}^{(i)} = W_{j,c}^{(i)} = \frac{1}{2(n + \lambda)} (j = 1, 2, \dots, 2n) \quad (12)$$

where  $\alpha, \beta$ , and  $\lambda$  are the coefficients in UKF,  $n$  is the number of state variables. On this basis, the state estimation of UKF can be written as follows:

$$r_{j,k|k-1}^{(i)} = r_{j,k-1}^{(i)} + u_{k-1}^{(i)} \quad (13)$$

$$r_{k|k-1}^{(i)} = \sum_{j=0}^{2n} W_{j,c}^{(i)} r_{j,k|k-1}^{(i)} \quad (14)$$

Furthermore, the expression for covariance estimation can be shown as follows:

$$S_{k|k-1}^{(i)} = \sum_{j=0}^{2n} \left[ W_{j,c}^{(i)} \left( r_{k|k-1}^{(i)} - r_{j,k|k-1}^{(i)} \right) \left( r_{k|k-1}^{(i)} - r_{j,k|k-1}^{(i)} \right)^T \right] + Q_k^{(i)} \quad (15)$$

where  $Q_k = \text{diag}(\sigma_{a_k}^2, \sigma_{b_k}^2, \sigma_{c_k}^2, \sigma_{d_k}^2)$ . Correspondingly, the forward one-step prediction is shown as follows:

$$y_{j,k|k-1}^{(i)} = h \left( r_{j,k-1}^{(i)} \right) + v_{k-1}^{(i)} \quad (16)$$

$$y_{k|k-1}^{(i)} = \sum_{j=0}^{2n} W_{j,m}^{(i)} y_{j,k|k-1}^{(i)} \quad (17)$$

On this basis, the self-covariance and cross-covariance are shown as follows:

$$S_{y_{k|k-1}, y_{k|k-1}}^{(i)} = \sum_{j=0}^{2n} W_{j,c}^{(i)} \left( y_{k|k-1}^{(i)} - y_{j,k|k-1}^{(i)} \right)^2 + (\sigma_{v_k}^2)^{(i)} \quad (18)$$

$$S_{r_{k|k-1}, y_{k|k-1}}^{(i)} = \sum_{j=0}^{2n} \left[ W_{j,c}^{(i)} \left( r_{k|k-1}^{(i)} - r_{j,k|k-1}^{(i)} \right) \left( y_{k|k-1}^{(i)} - y_{j,k|k-1}^{(i)} \right) \right] \quad (19)$$

In this sense, the UKF gain  $K_k$  is shown as follows:

$$K_k^{(i)} = \left( S_{y_{k|k-1}, y_{k|k-1}}^{(i)} \right)^{-1} S_{r_{k|k-1}, y_{k|k-1}}^{(i)} \quad (20)$$

Through the addition of new battery capacity data during the  $k$ th operational process, the state and the covariance update of UKF can be expressed as follows:

$$r_k^{(i)} = r_{k|k-1}^{(i)} + K_k^{(i)} (y_k - y_{k|k-1}^{(i)}) \quad (21)$$

$$S_k^{(i)} = S_{k|k-1}^{(i)} + K_k^{(i)} S_{y_{k|k-1}, y_{k|k-1}}^{(i)} \left( K_k^{(i)} \right)^T. \quad (22)$$

According to (21)–(22), the posterior PDF of particles under UKF algorithm can be obtained.

3) The weight update of the particles:

According to step 2, the importance function of the  $i$ th particle can be obtained as:  $\hat{r}_k^{(i)} \sim p(r_k^{(i)} | r_{0:k-1}^{(i)}, y_{1:k})$ . Consequently, the weight update expression of the particles can be written as follows:

$$w_k^{(i)} = w_{k-1}^{(i)} \frac{p(y_k | \hat{r}_k^{(i)}) p(\hat{r}_k^{(i)} | r_{k-1}^{(i)})}{p(\hat{r}_k^{(i)} | r_{0:k-1}^{(i)}, y_{1:k})} \quad (23)$$

where  $p(y_k | \hat{r}_k^{(i)})$  and  $p(\hat{r}_k^{(i)} | r_{k-1}^{(i)})$  can be calculated by state transition equation and measurement equation. The normalized particle weighted can be obtained as follows:

$$w_{kn}^{(i)} = \frac{w_k^{(i)}}{\sum_{i=1}^{N_s} w_k^{(i)}}. \quad (24)$$

4) Particle resampling:

To promote the state estimation performance of the UPF algorithm, this article adopts the resampling method to retain the particles with larger weights, and to reduce the particles with smaller weights. It is assumed that the number of effective particles is  $N_{ef}$ , which can be defined as follows:

$$N_{ef} = \frac{1}{\sum_{i=1}^{N_s} (w_k^{(i)})^2}. \quad (25)$$

The threshold is set to  $N_{th}$ . When the number of effective particles  $N_{ef}$  is lower than  $N_{th}$ , particle resampling is required. The  $N_s$  particles  $\tilde{r}_k^{(i)}$  are sampled from particle set  $\hat{r}_k$ , and the weight of the  $i$ th particle is  $\tilde{w}_{kn}^{(i)}$ , shown as follows:

$$\tilde{w}_{kn}^{(i)} = \frac{1}{N_s}. \quad (26)$$

5) State and covariance update:

On the basis of the particle weights obtained by resampling, the state variables and covariances under the posterior conditions of the UPF algorithm can be acquired as follows:

$$\hat{r}_{k|k} = \sum_{i=1}^{N_s} \tilde{w}_{kn}^{(i)} \hat{r}_k^{(i)} \quad (27)$$

$$\hat{S}_{k|k} = \sum_{i=1}^{N_s} \tilde{w}_{kn}^{(i)} \left( \hat{r}_k^{(i)} - \hat{r}_{k|k} \right) \left( \hat{r}_k^{(i)} - \hat{r}_{k|k} \right)^T. \quad (28)$$

## B. Adaptive EM Parameter Estimation Algorithm

Regarding UPF degradation model, the variances of noise in the state transition equation and measurement equation are unknown. Hence, the battery capacity data in the degradation process are considered, and the EM algorithm is introduced to perform adaptive parameter estimation for unknown parameters [35]. Specifically, the unknown parameter vector is written as:  $\Xi = [\sigma_a^2, \sigma_b^2, \sigma_c^2, \sigma_d^2, \sigma_v^2]^T$ . The joint likelihood function

$p(x_{0:k}, y_{1:k} | \Xi)$  is maximized to accomplish the estimation of unknown parameters in  $\Xi$ . In this sense, the joint log-likelihood function is written as follows:

$$\begin{aligned} L_k(\Xi) &= \sum_{i=1}^{N_s} \ln p \left( r_{1:k}^{(i)}, y_{1:k} | \Xi \right) \\ &= \sum_{i=1}^{N_s} \ln \left[ p \left( r_{1:k}^{(i)} | \Xi \right) p \left( y_{1:k} | r_{1:k}^{(i)}, \Xi \right) \right] \\ &= \sum_{i=1}^{N_s} \left[ \sum_{t=1}^k \ln p \left( r_t^{(i)} | r_{t-1}^{(i)}, \Xi \right) + \sum_{t=1}^k \ln p \left( y_t | r_t^{(i)}, \Xi \right) \right] \end{aligned} \quad (29)$$

where  $r_t^{(i)} = [a_t^{(i)}, b_t^{(i)}, c_t^{(i)}, d_t^{(i)}]^T$ . On the basis of (4)–(6) in the adaptive UPF model, the relationship in (30)–(31) can be acquired

$$r_t^{(i)} | r_{t-1}^{(i)}, \Xi \sim N \left( r_{t-1}^{(i)}, U_t^{(i)} \right) \quad (30)$$

$$y_t | r_t^{(i)}, \Xi \sim N \left( a_t^{(i)} e^{b_t^{(i)} t} + c_t^{(i)} e^{d_t^{(i)} t}, (\sigma_{vt}^2)^{(i)} \right) \quad (31)$$

where  $U_t^{(i)} = [(\sigma_{at}^2)^{(i)}, (\sigma_{bt}^2)^{(i)}, (\sigma_{ct}^2)^{(i)}, (\sigma_{dt}^2)^{(i)}]^T$ . Substituting (30)–(31) into (29), the joint log-likelihood function can be further expressed as follows:

$$\begin{aligned} L_k(\Xi) &= \sum_{i=1}^{N_s} \left[ \sum_{t=1}^k \ln p \left( r_t^{(i)} | r_{t-1}^{(i)}, \Xi \right) + \sum_{t=1}^k \ln p \left( y_t | r_t^{(i)}, \Xi \right) \right] \\ &\propto \sum_{i=1}^{N_s} \left\{ -\frac{1}{2} \sum_{t=1}^k \left[ \ln |Q_t^{(i)}| + \left( r_t^{(i)} - r_{t-1}^{(i)} \right)^T \left( Q_t^{(i)} \right)^{-1} \right. \right. \\ &\quad \times \left. \left( r_t^{(i)} - r_{t-1}^{(i)} \right) \right] \\ &\quad \left. - \frac{1}{2} \sum_{t=1}^k \left[ \ln (\sigma_{vt}^2)^{(i)} + \frac{\left( y_t - a_t^{(i)} e^{b_t^{(i)} t} + c_t^{(i)} e^{d_t^{(i)} t} \right)^2}{(\sigma_{vt}^2)^{(i)}} \right] \right\}. \end{aligned} \quad (32)$$

For the  $m$ th iteration process in EM algorithm, the unknown parameter estimation vector can be written as:  $\hat{\Xi}_k^{(m)} = [(\sigma_{ak}^2)^{(m)}, (\sigma_{bk}^2)^{(m)}, (\sigma_{ck}^2)^{(m)}, (\sigma_{dk}^2)^{(m)}, (\sigma_{vk}^2)^{(m)}]^T$ . According to the derivation of (30)–(32), and the basic principle of EM algorithm, the  $m+1$ th iteration process consists of E step and M step, which can be expressed as follows:

1) E step:

$$L_k \left( \Xi | \hat{\Xi}_k^{(m)} \right) = E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}} [L_k(\Xi)] \quad (33)$$

2) M step:

$$\hat{\Xi}_k^{(m+1)} = \arg \max_{\Xi} \left\{ E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}} [L_k(\Xi)] \right\} \quad (34)$$

According to (32), a joint PDF of the unknown parameter vector  $\Xi$  is specified, which is a key for the E step, shown in (35) shown at the bottom of next page.



In (35),  $E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}}(r_t^{(i)})$ ,  $E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}}(r_t^{(i)}(r_t^{(i)})^T)$ ,  $E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}}(r_t^{(i)}(r_{t-1}^{(i)})^T)$ , and  $E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}}(r_{t-1}^{(i)}(r_t^{(i)})^T)$  are the conditional expectations based on the state variables of the battery capacity data  $y_{1:k} = [y_1, y_2, \dots, y_k]$ . To calculate these expectations under the posterior condition, the Rauch–Tung–Striebel (RTS) algorithm is implemented for backward smoothing [36], [37].

According to the forward iteration of the UPF algorithm, the estimations of the expectation and covariance of the state variables represented by the particles can be shown as follows:

$$\tilde{r}_{k|k}^{(i)} = \hat{r}_{k|k}^{(i)} \quad (36)$$

$$\tilde{S}_{k|k}^{(i)} = \hat{S}_{k|k}^{(i)}. \quad (37)$$

Correspondingly, the covariance of state variables between the  $k$ -1th and the  $k$ th cycles can be shown as follows:

$$\tilde{S}_{k-1,k|k}^{(i)} = \text{cov}(r_{k-1}^{(i)}, r_k^{(i)} | y_{1:k}) = (I_{4 \times 4} - K_k^{(i)}) \hat{S}_{k-1|k-1}^{(i)}. \quad (38)$$

The RTS smoothing can be implemented based on the initialization in (36)–(38). The RTS smoothing gain can be calculated as follows:

$$G_t^{(i)} = S_{t|t}^{(i)} \left( S_{t+1|t}^{(i)} \right)^{-1}. \quad (39)$$

Correspondingly, the update of the state variables and covariances in the backward iteration process are shown as follows:

$$\tilde{r}_{t|k}^{(i)} = \tilde{r}_{t|t}^{(i)} + G_t^{(i)} \left( \tilde{r}_{t+1|k}^{(i)} - \hat{r}_{t+1|t}^{(i)} \right) \quad (40)$$

$$\tilde{S}_{t|k}^{(i)} = S_{t|t}^{(i)} + G_t^{(i)} \left( \tilde{S}_{t+1|k}^{(i)} - S_{t+1|t}^{(i)} \right) \left( G_t^{(i)} \right)^T. \quad (41)$$

Moreover, the covariance between both neighboring cycles of state variables can be shown as follows:

$$\tilde{S}_{t-1,t|k}^{(i)} = S_{t|t}^{(i)} \left( G_{t-1}^{(i)} \right)^T + G_t^{(i)} \left( \tilde{S}_{t,t+1|k}^{(i)} - S_{t|t}^{(i)} \right) \left( G_{t-1}^{(i)} \right)^T. \quad (42)$$

On the basis of (36)–(42), the conditional expectations are written as follows:

$$E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}}(r_t^{(i)}) = \tilde{r}_{t|k}^{(i)} \quad (43)$$

$$E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}}(r_t^{(i)}(r_t^{(i)})^T) = \tilde{r}_{t|k}^{(i)} \left( \tilde{r}_{t|k}^{(i)} \right)^T + \tilde{S}_{t|k}^{(i)} \quad (44)$$

$$E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}}(r_t^{(i)}(r_{t-1}^{(i)})^T) = \tilde{r}_{t|k}^{(i)} \left( \tilde{r}_{t-1|k}^{(i)} \right)^T + \tilde{S}_{t-1,t|k}^{(i)} \quad (45)$$

$$E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}}(r_{t-1}^{(i)}(r_t^{(i)})^T) = \tilde{r}_{t-1|k}^{(i)} \left( \tilde{r}_{t|k}^{(i)} \right)^T + \tilde{S}_{t-1,t|k}^{(i)}. \quad (46)$$

The conditional expectations in (43)–(46) are substituted into (35), the extended form of which can be written as (47) shown at the bottom of next page. Based on the calculation results in (36)–(47) of the M step can be acquired through (34). Regarding the unknown vector  $\hat{\Xi}_k$  in UPF model, the partial derivative of each element is set to 0. On this basis, the results of (48)–(49) is solved.

$$\begin{aligned} Q_k^{(m+1)} = \frac{1}{kN_s} \sum_{i=1}^{N_s} \sum_{t=1}^k & \left[ \left( \tilde{r}_{t|k}^{(i)} \left( \tilde{r}_{t|k}^{(i)} \right)^T + \tilde{S}_{t|k}^{(i)} \right) \right. \\ & - \left( \tilde{r}_{t|k}^{(i)} \left( \tilde{r}_{t-1|k}^{(i)} \right)^T + \tilde{S}_{t-1,t|k}^{(i)} \right) \\ & - \left( \tilde{r}_{t-1|k}^{(i)} \left( \tilde{r}_{t|k}^{(i)} \right)^T + \tilde{S}_{t-1,t|k}^{(i)} \right) \\ & \left. + \left( \tilde{r}_{t-1|k}^{(i)} \left( \tilde{r}_{t-1|k}^{(i)} \right)^T + \tilde{S}_{t-1|k}^{(i)} \right) \right] \end{aligned} \quad (48)$$

$$(\sigma_{vk}^2)^{(m+1)} = \frac{1}{kN_s} \sum_{i=1}^{N_s} \sum_{t=1}^k \left[ \left( y_t - a_t^{(i)} e^{b_t^{(i)} \cdot t} + c_t^{(i)} e^{d_t^{(i)} \cdot t} \right)^2 \right] \quad (49)$$

where  $Q_k^{(m+1)} = \text{diag}[(\sigma_{ak}^2)^{(m+1)}, (\sigma_{bk}^2)^{(m+1)}, (\sigma_{ck}^2)^{(m+1)}, (\sigma_{dk}^2)^{(m+1)}]$ . By continuously performing the iteration of

$$\begin{aligned} L_k(\Xi | \hat{\Xi}_k^{(m)}) &= E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}}[L_k(\Xi)] \\ &\propto E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}} \left\{ \sum_{i=1}^{N_s} \left\{ -\frac{1}{2} \sum_{t=1}^k \left[ \ln |Q_k| + (r_t^{(i)} - r_{t-1}^{(i)}) Q_k^{-1} (r_t^{(i)} - r_{t-1}^{(i)})^T \right] - \frac{1}{2} \sum_{t=1}^k \left[ \ln \sigma_v^2 + \frac{(y_t - a_t^{(i)} e^{b_t^{(i)} \cdot t} + c_t^{(i)} e^{d_t^{(i)} \cdot t})^2}{\sigma_v^2} \right] \right\} \right\} \\ &\propto N_s \left( -\frac{k}{2} \ln |Q_k^{(m)}| - \frac{k}{2} \ln (\sigma_{vk}^2)^{(m)} \right) + \sum_{i=1}^{N_s} \left\{ -\frac{1}{2} \sum_{t=1}^k \left\{ \left( Q_k^{(m)} \right)^{-1} E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}} \left[ r_t^{(i)} (r_t^{(i)})^T - r_{t-1}^{(i)} (r_{t-1}^{(i)})^T \right. \right. \right. \\ &\quad \left. \left. - r_{t-1}^{(i)} (r_t^{(i)})^T + r_{t-1}^{(i)} (r_{t-1}^{(i)})^T \right] \right\} \\ &\quad \left. - \frac{1}{2 (\sigma_{vk}^2)^{(m)}} \sum_{t=1}^k E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}} \left[ \left( y_t - a_t^{(i)} e^{b_t^{(i)} \cdot t} + c_t^{(i)} e^{d_t^{(i)} \cdot t} \right)^2 \right] \right\} \end{aligned} \quad (35)$$

the E step or the M step, only the iteration is stopped when the maximum number of iterations reaches 10 times or  $\|\hat{\Xi}_k^{(m+1)} - \hat{\Xi}_k^{(m)}\|_1 \leq \varepsilon$  is satisfied. On this basis, the adaptive estimation of the noise in the state transition equation and measurement equation is accomplished.

### C. Capacity Regeneration Point Detection Based on UPF-W

On the basis of adaptive UPF model, the posterior PDF of particles in the UKF can be obtained through (7)–(22), which is used as the prior PDF of the PF algorithm. In addition, the posterior PDF of PF can be obtained through (23)–(28). For the convenience of expression, the battery capacity of the particles through UKF algorithm is  $\hat{q}_k = \{\hat{q}_k^{(1)}, \hat{q}_k^{(2)}, \dots, \hat{q}_k^{(N_s)}\}$ . The battery capacity after the update through the PF algorithm is  $\tilde{q}_k = \{\tilde{q}_k^{(1)}, \tilde{q}_k^{(2)}, \dots, \tilde{q}_k^{(N_s)}\}$ . Whereas for the capacity regeneration phenomenon after the battery standing for a period of time, a certain distribution difference is existed between  $\hat{q}_k$  and  $\tilde{q}_k$ .

The Wilcoxon rank sum test is a nonparametric test method, the contribution of which is to measure the distribution difference between two groups of data samples [38]. Without any special assumptions about the distribution of objective data, the Wilcoxon rank sum test can be applied to some complicated distribution situations. Consequently, the Wilcoxon rank sum test is used to measure the distribution difference between  $\hat{q}_k$  and  $\tilde{q}_k$ . The hypothesis  $H_0$  is proposed as follows:

- 1) The two populations  $\hat{q}_k$  and  $\tilde{q}_k$  have the same distribution at the significant level  $\gamma$ .

To verify the hypothesis  $H_0$ , the Wilcoxon rank sum test can be summarized as the Algorithm 1. To sum up, once the hypothesis  $H_0$  is rejected, it means there is a big difference in the distribution of  $\hat{q}_k$  and  $\tilde{q}_k$ . In this sense, the battery has a capacity regeneration phenomenon in the  $k$ th operational cycle.

### D. Summary of RUL Prediction Approach

Based on the adaptive parameter estimation of the EM algorithm in Section III-B, the noise variables of the state transition equation and measurement equation in the battery degradation model can be calculated. In this sense, the trend extrapolation is performed by the updated values of state variables in the

#### Algorithm 1: Wilcoxon Rank Sum Test.

**Input:** The two populations  $\hat{q}_k$  and  $\tilde{q}_k$ .

**Output:** Accept or reject hypothesis  $H_0$ .

- 1: The  $Q_k$  is obtained by mixing the  $\hat{q}_k$  and the  $\tilde{q}_k$ .
- 2: The elements in  $Q_k$  are numbered according to the value from small to large, the ordinal number corresponding to each sample is defined as the rank.
- 3: Calculate the sum of ranks  $T$  in the population  $\hat{q}_k$ .
- 4: According to  $N_s$  and  $\gamma$ , the rank sum test table is consulted, so that rank sum lower limit  $T_1$ , the rank sum upper limit  $T_2$  and the confidence level  $p$  can be obtained.
- 5: Set the interval  $[T_1, T_2]$ , if  $T$  is inside the interval,  $p > \gamma$ ,  $H_0$  is accepted. Alternatively,  $H_0$  is rejected.

capacity degradation model obtained in Section III-A. Moreover, in Section III-C, the UPF-W algorithm is proposed to detect capacity regeneration points to reduce the influence of capacity regeneration on the RUL prediction. The predicted RUL of the battery is the number of cycles that elapsed from current moment to the time when the capacity first degrades to the failure threshold. Finally, by substituting each particle into the model, the confidence interval of the capacity can be acquired to measure the uncertainty of the battery capacity during the degradation process.

To sum up, the overall flow of EM-UPF-W-based RUL prediction approach can be summarized in Fig. 1, where SP denotes start point, CRP denotes capacity regeneration point. Furthermore, the proposed EM-UPF-W RUL prediction approach can be summarized in Algorithm 2.

## IV. CASE STUDY

### A. Dataset Description

A battery dataset produced by NASA is used to demonstrate the EM-UPF-W-based RUL prediction approach. The dataset includes the data generated by four 18 650 lithium-ion batteries, which are numbered B0005, B0006, B0007, and B0018. The operation of the four batteries with a rated capacity of 2Ah include a charging mode and a discharging mode. For the charging mode, the battery is charged with a constant current until the battery

$$\begin{aligned}
 L_k(\Xi | \hat{\Xi}_k^{(m)}) &= E_{r_{1:k}, y_{1:k} | \hat{\Xi}_k^{(m)}} [L_k(\Xi)] \\
 &\propto N_s \left( -\frac{k}{2} \ln |Q_k^{(m)}| - \frac{k}{2} \ln (\sigma_{vk}^{(m)})^2 \right) - \sum_{i=1}^{N_s} \left\{ \frac{1}{2} \sum_{t=1}^k (Q_k^{(m)})^{-1} \left[ \left( \tilde{r}_{t|k}^{(i)} (\tilde{r}_{t|k}^{(i)})^T + \tilde{S}_{t|k}^{(i)} \right) - \left( \tilde{r}_{t|k}^{(i)} (\tilde{r}_{t-1|k}^{(i)})^T + \tilde{S}_{t-1,t|k}^{(i)} \right) \right. \right. \\
 &\quad \left. \left. - \left( \tilde{r}_{t|k}^{(i)} (\tilde{r}_{t-1|k}^{(i)})^T + \tilde{S}_{t-1,t|k}^{(i)} \right) \right] \right. \\
 &\quad \left. + \left( \tilde{r}_{t-1|k}^{(i)} (\tilde{r}_{t-1|k}^{(i)})^T + \tilde{S}_{t-1|k}^{(i)} \right) \right] + \frac{1}{2 (\sigma_{vk}^{(m)})^2} \sum_{t=1}^k \left[ \left( y_t - a_t^{(i)} e_t^{(i) \cdot t} + c_t^{(i)} e_{d_t^{(i)} \cdot t} \right)^2 \right] \right\}
 \end{aligned} \tag{47}$$

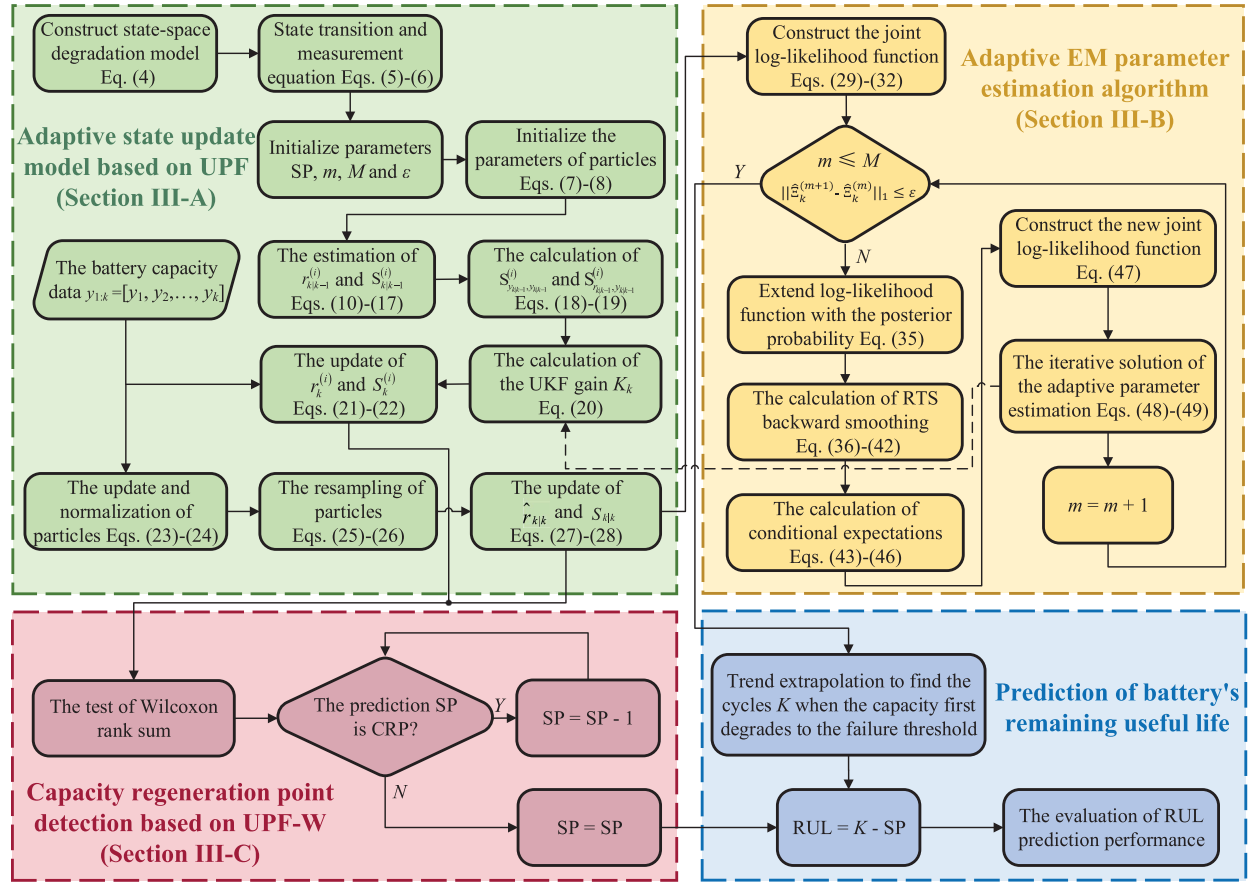


Fig. 1. Overall flow diagram of EM-UPF-W approach.

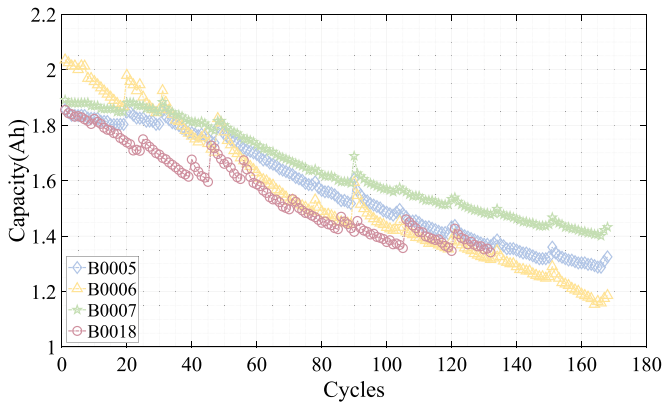


Fig. 2. Battery capacity decay curves.

voltage increases to 4.2 V, and then the constant voltage charging is switched until the charging current drops to 20 mA. For the discharging mode, these batteries are discharged at a 2 A current until the voltages drop to different thresholds, respectively. The battery capacity decay curves of these four batteries are shown in Fig. 2. Considering the difference in the operational conditions of the four batteries, the failure thresholds are 1.4Ah, 1.22Ah, 1.6Ah, and 1.4Ah, respectively [12], [29].

## B. Implementation Details of the Battery Degradation Model

In order to obtain a relatively suitable initial vector value, when the RUL prediction of one battery is performed, the double exponential function fitting coefficients' mean of the battery capacity degradation curves of the other three batteries are used as the initial state variable value. For instance, when predicting the RUL of battery B0005, the double exponential function fitting coefficients' mean of the capacity degradation curves of batteries B0006, B0007, and B0018, are used as the value of the initial state variable. On this basis, the initial values of these batteries are shown in Table I.

Furthermore, the influence of the initial variable values on RUL prediction is discussed in this article. Since the variable  $c$  is more sensitive to capacity estimation, the variable  $c$  is used as an example to discuss the impact of different parameter initial values on RUL prediction. The initial value of  $c$  is to  $-0.5$ ,  $0$ ,  $0.16$ , and  $0.5$ , respectively. It should be noted that the values of other parameters are consistent. Taking the 60th operational cycle as the prediction start point, Fig. 3 shows the influence of these four different initial values on RUL prediction. The green part and blue part denote the 99% confidence interval of the capacity and the failure cycle. As can be seen from Fig. 3, the first few operational cycles have deviation between the predicted

**Algorithm 2:** EM-UPF-W-Based RUL Prediction Algorithm.

**Input:** Capacity data  $y_{1:k} = [y_1, y_2, \dots, y_k]$ , maximum number of iterations  $M$ , error threshold  $\varepsilon$

**Output:** The  $RUL_k$  under the  $k$ th operational cycle and the corresponding confidence interval

- 1: Construct battery capacity degradation model based on state space according to Eqs. (4)-(6).
- 2: Initialize the particles and start point  $SP = k$ .
- 3: UKF is updated to obtain  $\hat{S}_k$  and  $\hat{r}_k$  based on Eqs. (9)-(22).
- 4: UPF is updated to obtain  $\hat{S}_{k|k}$  and  $\hat{r}_{k|k}$  by Eqs. (23)-(28).
- 5: Establish joint likelihood function through Eqs. (29)-(32).
- 6: **while**  $m \leq M$  or  $\|\hat{\Xi}_k^{(m+1)} - \hat{\Xi}_k^{(m)}\|_1 \leq \varepsilon$  **do**
- 7:   Extend log-likelihood function under the posterior probability based on Eq. (35).
- 8:   Calculate conditional expectations in Eqs. (43)-(46) through RTS backward smoothing by Eqs. (36)-(42).
- 9:   The partial derivative of each element in the unknown vector  $\hat{\Xi}_k$  in UPF model is set to 0, so as to acquire the updated results by Eqs. (47)-(49).
- 10:    $m \rightarrow m + 1$ .
- 11: **end while**
- 12: Determine whether CRP occurs according to Algorithm 1.
- 13: **if** CRP occurs **then**
- 14:    $SP = SP - 1$
- 15: **else**
- 16:    $SP = SP$
- 17: **end if**
- 18: Trend extrapolation to find the cycles  $K$  when the capacity first degrades to the failure threshold.
- 19: Calculate  $RUL_k = K - SP$
- 20: Acquire the confidence interval of the battery's RUL by each particle in UPF model.

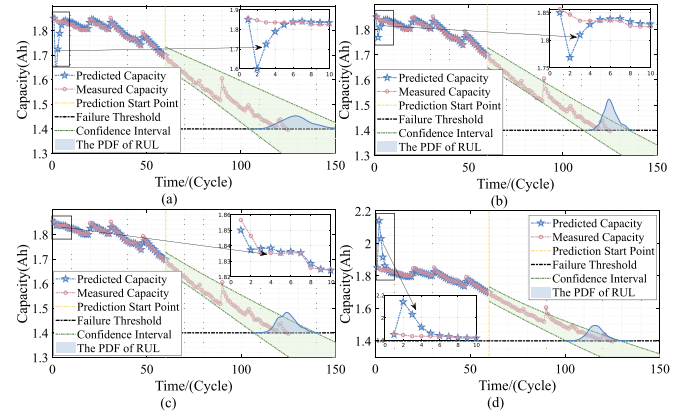
capacity and measured capacity. However, with the addition of more capacity data, the predicted capacity of the UPF model will converge to the measured capacity, which represents the great adaptability of the proposed approach in this article. Moreover, the predicted RUL values are basically the same, even if there are differences in the initial values.

### C. Adaptive Noise Estimation Results

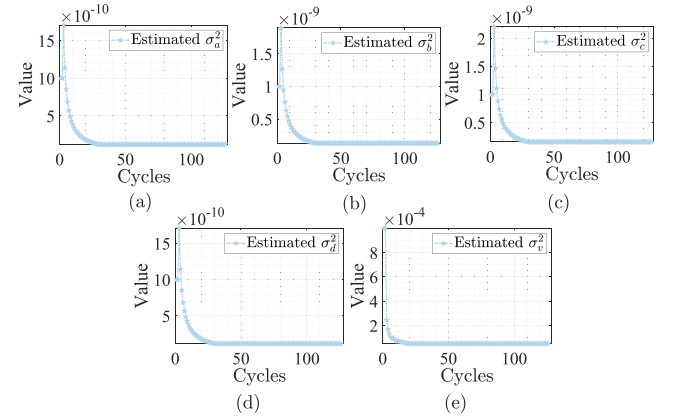
Only the capacity data during the degradation process are considered in the EM algorithm. In this sense, the noise variables of the state transition equation and measurement equation based on the UPF model can be acquired. The B0005 battery is taken as an example to represent the experimental results of adaptive noise estimation based on EM algorithm. Specifically, for the initial unknown parameter vector,  $\hat{\Xi}_0 = [\sigma_{a0}^2, \sigma_{b0}^2, \sigma_{c0}^2, \sigma_{d0}^2, \sigma_{v0}^2]^T$ , the initial value,  $\sigma_{a0}^2 = \sigma_{b0}^2 = \sigma_{c0}^2 = \sigma_{d0}^2 = 10^{-9}$ , the initial value,  $\sigma_{v0}^2 = 10^{-3}$ . With the aid of the operational capacity data, the

**TABLE I**  
INITIAL VALUES OF THE FOUR BATTERIES

| Battery | The double exponential fitting coefficients |         |         |          |
|---------|---|---------|---------|----------|
|         | $a$   | $b$     | $c$     | $d$      |
| B0005   | 1.98  | -0.0027 | -0.17   | -0.069   |
| B0006   | 1.57  | -0.0056 | 0.49    | 9.45e-04 |
| B0007   | 1.94  | -0.0021 | 1.57e-7 | 0.074    |
| B0018   | 1.86  | -0.0029 | 1.91e-4 | 0.048    |
| Battery | The initial value of state variables        |         |         |          |
|         | $a_0$                                       | $b_0$   | $c_0$   | $d_0$    |
| B0005   | 1.79  | -0.0035 | 0.16    | 0.041    |
| B0006   | 1.93  | -0.0026 | -0.057  | 0.018    |
| B0007   | 1.80  | -0.0037 | 0.11    | -0.0067  |
| B0018   | 1.83  | -0.0035 | 0.11    | 0.0020   |



**Fig. 3.** Influence of the four different initial values on RUL prediction. (a)  $c = -0.5$ . (b)  $c = 0$ . (c)  $c = 0.16$ . (d)  $c = 0.5$ .

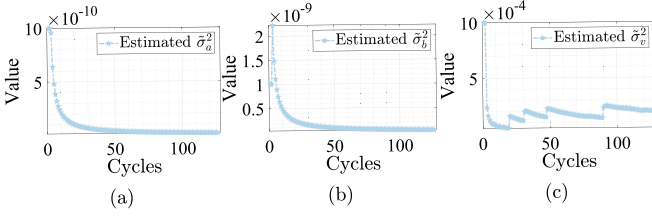


**Fig. 4.** Adaptive noise estimation results through EM-UPF-based noise updating algorithm. (a) The estimation of  $\sigma_a^2$ . (b) The estimation of  $\sigma_b^2$ . (c) The estimation of  $\sigma_c^2$ . (d) The estimation of  $\sigma_d^2$ . (e) The estimation of  $\sigma_v^2$ .

model parameters can be adaptively updated. Fig. 4 represents the noise estimation results of the B0005 battery over the full life cycle.

To demonstrate the superiority of the proposed EM-UPF approach, this article reproduces the EM-KF-based adaptive noise updating algorithm in [12] for B0005 battery. Considering that the KF algorithm is oriented to linear systems, so the linear





**Fig. 5.** Adaptive noise estimation results through EM-KF-based noise updating algorithm. (a) The estimation of  $\sigma_a^2$ . (b) The estimation of  $\sigma_b^2$ . (c) The estimation of  $\sigma_v^2$ .

battery degradation model can be written as (50)

$$\begin{cases} \tilde{x}_k = \tilde{x}_{k-1} + \tilde{u}_k \\ \tilde{y}_k = \tilde{h}(\tilde{x}_k) + \tilde{v}_k \end{cases} \quad (50)$$

where  $\tilde{x}_k = [\tilde{a}_k, \tilde{b}_k]^T$ ,  $\tilde{u}_k = [\tilde{u}_a, \tilde{u}_b]^T$ ,  $\tilde{v}_k \in R^{1 \times 1}$ . The state transition in (50) can be written in (51)

$$\begin{cases} \tilde{a}_k = \tilde{a}_{k-1} + \tilde{u}_a & (\tilde{u}_a \sim N(0, \sigma_a^2)) \\ \tilde{b}_k = \tilde{b}_{k-1} + \tilde{u}_b & (\tilde{u}_b \sim N(0, \sigma_b^2)) \end{cases} \quad (51)$$

where  $\tilde{u}_a, \tilde{u}_b$  are the noise terms for state variables, respectively,  $\sigma_a^2, \sigma_b^2$  are the variances of the corresponding noise terms, respectively. Correspondingly, the measurement equation can be shown in (52)

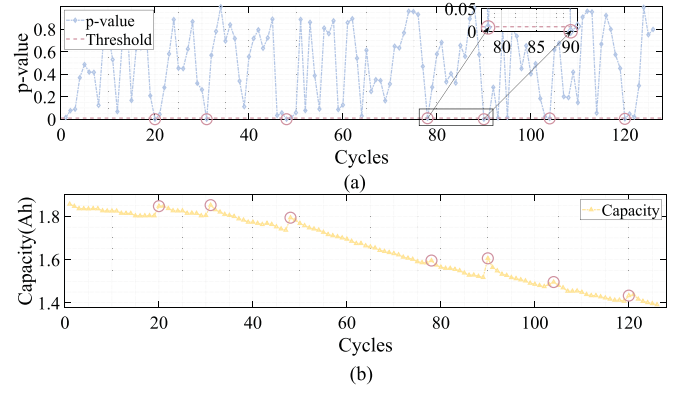
$$\tilde{y}_k = \tilde{a}_k + \tilde{b}_k \cdot k \quad (52)$$

where  $\tilde{v}_k$  is the measured noise, the variance of which is  $\sigma_v^2$ . The adaptive noise estimation result is shown in Fig. 5. Compared with Figs. 4 and 5, it can be seen that EM-UPF-based noise updating algorithm has faster convergence speed than EM-KF-based noise updating algorithm in [12]. In detail, all the unknown parameters can converge within 20 operational cycles. Furthermore, EM-UPF-based noise updating algorithm can make the measurement noise converge as quickly as possible. In this sense, the adaptive parameter estimation approach based on the EM-UPF algorithm has a great convergence performance.

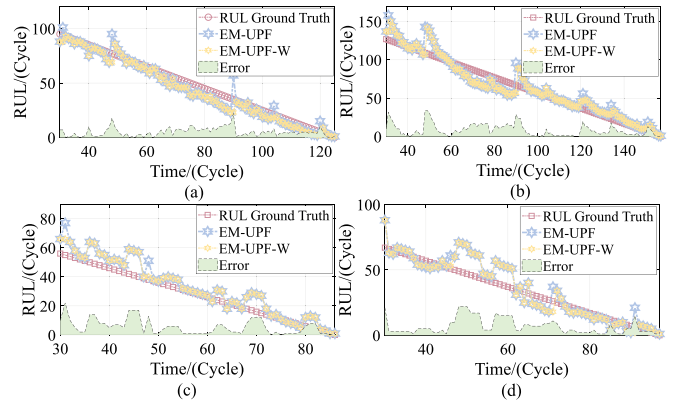
#### D. Capacity Regeneration Point Detection Results Based on UPF-W

Fig. 2 represents that the overall battery capacity decreases gradually during the operation. Nevertheless, during the battery's degradation process, there are some capacity regeneration points, the capacity of which tends to abruptly change. If these capacity regeneration points are directly used as the starting point of RUL prediction, there will be a large predicted error. The reason is that the predicted RUL value is higher than the actual value. Consequently, the UPF-W algorithm is proposed to eliminate the effect.

Taking B0005 battery as an example, the battery has obvious capacity regeneration points in the 20th, 31st, 48th, 78th, 90th, 104th, and 120th cycles. In this article, the number of particles,  $N_s$  is 500, the confidence interval of the Wilcoxon rank sum test is 0.01. Correspondingly, according to Section III-A, the battery capacity of particles after being updated by the UKF



**Fig. 6.** Detection results of the capacity regeneration points. (a) The experimental results of Wilcoxon ranksum test. (b) The decay curve of the battery capacity.



**Fig. 7.** RUL prediction results based on EM-UPF and EM-UPF-W. (a) B0005 Battery. (b) B0006 Battery. (c) B0007 Battery. (d) B0018 Battery.

algorithm and the PF algorithm are  $\hat{q}_k = \{\hat{q}_k^{(1)}, \hat{q}_k^{(2)}, \dots, \hat{q}_k^{(N_s)}\}$ ,  $\tilde{q}_k = \{\tilde{q}_k^{(1)}, \tilde{q}_k^{(2)}, \dots, \tilde{q}_k^{(N_s)}\}$ , respectively.

Fig. 6 represents the detection results of the UPF-W algorithm for the capacity regeneration point for the B0005 battery from the initial state to the failure. In detail, the sub-figure (a) is the result of the  $p$ -value of the Wilcoxon rank sum test, and the sub-figure (b) is the corresponding decay curve of the battery capacity. For the convenience of expression, the capacity regeneration points of the batteries are marked with red circles. Through Fig. 6, it can be known that the capacity regeneration point detection approach based on UPF-W proposed in this article have a great detection performance.

#### E. Experimental Results Based on EM-UPF-W Approach

Fig. 7 represents the RUL predicted values of EM-UPF and EM-UPF-W, the RUL prediction error of EM-UPF-W, and the RUL ground truth for the four batteries. In detail, the blue line and yellow line represent the RUL prediction results of EM-UPF-W and EM-UPF, respectively. The red line and the green part represent the RUL ground truth and the RUL prediction error, respectively. Since the proposed EM-UPF-W is an adaptive

TABLE II  
COMPARISON OF EXPERIMENTAL RESULTS OF MULTIPLE EXISTING APPROACHES

| Method          | MAE (Cycle)  |              |              |              |              | RMSE (Cycle) |               |              |              |              |
|-----------------|--------------|--------------|--------------|--------------|--------------|--------------|---------------|--------------|--------------|--------------|
|                 | B0005        | B0006        | B0007        | B0018        | Average      | B0005        | B0006         | B0007        | B0018        | Average      |
| RFR             | 19.004       | 40.349       | 55.544       | 22.996       | 34.473       | 23.120       | 44.553        | 56.691       | 27.317       | 37.920       |
| ERT             | 17.796       | 40.279       | 55.935       | 21.976       | 33.997       | 23.004       | 44.735        | 57.360       | 27.275       | 38.094       |
| XGBoost         | 18.467       | 40.378       | 55.402       | 21.342       | 33.897       | 24.770       | 45.277        | 56.978       | 26.991       | 38.504       |
| GBT             | 16.826       | 38.875       | 55.767       | 20.128       | 32.900       | 19.467       | 42.709        | 56.111       | 22.722       | 35.252       |
| SVR             | 16.365       | 35.259       | 48.717       | 17.175       | 29.379       | 19.583       | 38.525        | 48.779       | 20.024       | 31.728       |
| RNN             | 9.196        | 35.483       | 10.647       | 16.956       | 18.071       | 12.401       | 38.369        | 12.452       | 19.818       | 20.760       |
| LSTM            | 11.974       | 31.177       | 16.850       | 10.421       | 17.606       | 15.379       | 35.980        | 20.594       | 11.680       | 20.908       |
| GRU             | 10.034       | 31.068       | 15.001       | 13.998       | 17.525       | 13.642       | 36.424        | 17.976       | 15.673       | 20.929       |
| KF [12]         | 9.358        | 13.681       | 9.324        | 10.769       | 10.783       | 10.563       | 15.879        | 10.687       | 12.219       | 12.337       |
| EKF [12]        | 8.325        | 12.687       | 8.754        | 9.316        | 9.771        | 9.658        | 14.503        | 9.758        | 11.872       | 11.448       |
| EM-KF-RTS [12]  | 7.337        | 10.333       | 7.263        | 8.716        | 8.412        | 8.684        | 13.869        | 8.002        | 11.215       | 10.443       |
| EM-UPF-L        | 7.066        | 10.045       | 6.909        | 8.412        | 8.108        | 8.267        | 13.572        | 8.098        | 10.885       | 10.206       |
| EM-UPF-SE       | 5.864        | 9.288        | 6.594        | 7.569        | 7.329        | 6.925        | 11.693        | 8.289        | 10.099       | 9.252        |
| EM-UPF          | 4.927        | 8.305        | 5.597        | 6.721        | 6.388        | 6.314        | 11.396        | 7.690        | 9.004        | 8.601        |
| <b>EM-UPF-W</b> | <b>4.583</b> | <b>7.508</b> | <b>5.210</b> | <b>6.382</b> | <b>5.921</b> | <b>5.653</b> | <b>10.100</b> | <b>7.062</b> | <b>8.695</b> | <b>7.878</b> |

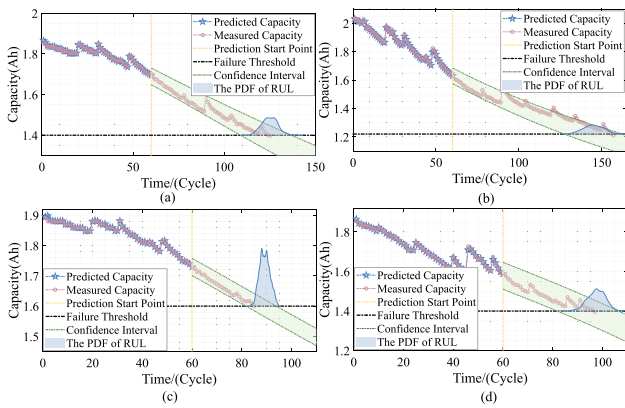


Fig. 8. Probability density interval of the capacity and failure cycle. (a) B0005 Battery. (b) B0006 Battery. (c) B0007 Battery. (d) B0018 Battery.

algorithm, the error of the battery's RUL gradually decreases with the increase of capacity data, which can be effectively confirmed in the error area represented by the green part in Fig. 7. The error in the later part of the battery life cycle is lower than that in the earlier part. Furthermore, it can be seen from Fig. 7 that the EM-UPF algorithm without the Wilcoxon rank sum test has a large error in the operational cycles of some capacity regeneration points. On the contrary, the proposed EM-UPF-W algorithm can acquire an excellent RUL prediction performance even on the cycles with battery capacity regeneration.

With an accurate RUL prediction of the battery, the proposed EM-UPF-W can also describe the uncertainty of the capacity and RUL during the battery degradation process. Specifically, by substituting each particle into a specific operational cycle, the capacity can be updated by EM-UPF. On this basis, the confidence interval of the capacity can be obtained by extrapolating the trends. In addition, the trend is extrapolated for each particle. The cycle that reaches the battery failure threshold for the first time during the extrapolation process can be regarded as the failure cycle of each particle. The battery's RUL represented by each particle can be obtained by subtracting the failure cycles from prediction start point. Therefore, according to all particles, the PDF interval of capacity and RUL can be obtained, the

TABLE III  
COMPARISON OF ABSOLUTE ERROR OF EXISTING APPROACHES ON SOME SPECIFIC OPERATIONAL CYCLES

| Cycle | Approach        | Absolute error (Cycle) |          |          |          |
|-------|-----------------|------------------------|----------|----------|----------|
|       |                 | B0005                  | B0006    | B0018    | Average  |
| 60    | PF [34]         | 39                     | 27       | 34       | 33       |
|       | PF-AR [34]      | 26                     | 24       | 14       | 21       |
|       | EKPF-AR [34]    | 18                     | 11       | 8        | 12       |
|       | <b>EM-UPF-W</b> | <b>7</b>               | <b>7</b> | <b>5</b> | <b>6</b> |
| 70    | PF [34]         | 33                     | 23       | 18       | 25       |
|       | PF-AR [34]      | 14                     | 17       | 5        | 12       |
|       | EKPF-AR [34]    | 9                      | 7        | 3        | 6        |
|       | <b>EM-UPF-W</b> | <b>1</b>               | <b>4</b> | <b>2</b> | <b>2</b> |

uncertainty of which can be well described. For instance, the 60th operational cycle of these four batteries is taken as an example, which is shown in Fig. 8. In detail, the green part and blue part denote the 99% confidence interval of the capacity and the failure cycle, respectively.

To demonstrate the superiority of the proposed EM-UPF-W approach, this article compares the approach with some existing learning-based approaches, for instance, support vector machine regression (SVR), random forest regression (RFR), extreme random tree (ERT), gradient boosting tree (GBT), extreme gradient boosting (XGBoost), recurrent neural network (RNN), long short-term memory network (LSTM), gated recurrent unit (GRU), and EM-UPF. It should be noted that the conventional learning-based approaches are supervised. In this article, each 18650 lithium-ion battery in the NASA dataset is used as the testing data to verify the algorithm model, and the other battery data are used as the historical data to train the algorithm model. In order to have a comparison with the existing hybrid empirical-based approaches in [12], KF, EKF, and EM-KF-RTS are reproduced using the battery thresholds in this article. Moreover, comparative experiments on the RUL prediction for various model forms, namely, linear model, EM-UPF-L ( $a_k + b_k \cdot k$ ), and single exponential model, EM-UPF-SE ( $a_k e^{b_k \cdot k}$ ) are used to demonstrate the effectiveness of the double exponential model for battery's RUL prediction. Taking the prediction results after the 30th operational cycle of the battery life as an example, Table II shows the MAE and RMSE of these approaches on the NASA battery dataset.

Furthermore, in order to compare the proposed RUL prediction approach with existing techniques better, some specific operational cycles are considered. Table III lists the comparison of the absolute error of some hybrid empirical-based approaches, namely, PF, PF-AR, EKPF-AR and the proposed EM-UPF-W approach. It can be seen from Tables II and III that, compared with the existing learning-based approaches, and some hybrid empirical-based approaches, the proposed EM-UPF-W-based RUL prediction approach can achieve great prediction performance.

## V. CONCLUSION

A novel EM-UPF-W approach is proposed to predict the RUL of the 18650 lithium-ion battery, which a vital component in the battery system of electric vehicles. In the case of unlabeled small samples, a dynamic degradation model is constructed with the aid of EM-UPF-W to describe the degradation process of a single battery. The main contributions of the proposed RUL prediction approach can be summarized as follows: 1) A dynamic degradation model based on UPF is proposed for a single battery, which can describe the uncertainty in the degradation process. Furthermore, the EM algorithm is introduced to adaptively estimate the noise variables of the state transition equation and measurement equation in the UPF model. 2) The UPF-W algorithm is proposed to detect the capacity regeneration point so that the large prediction error caused by the sudden change of battery capacity can be eliminated. 3) In this article, a battery dataset is used to demonstrate the effectiveness of the proposed EM-UPF-W approach. Experimental results represent that the approach outperforms some existing data-driven methods.

## REFERENCES

- [1] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques-part I: Fault diagnosis with model-based and signal-based approaches," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3757–3767, Jun. 2015.
- [2] Z. Gao and S. Sheng, "Real-time monitoring, prognosis, and resilient control for wind turbine systems," *Renewable Energy*, vol. 116, pp. 1–4, 2018.
- [3] T. Jiang et al., "Development of a decentralized smart charge controller for electric vehicles," *Int. J. Elect. Power Energy Syst.*, vol. 61, pp. 355–370, 2014.
- [4] S. Ansari et al., "Multi-channel profile based artificial neural network approach for remaining useful life prediction of electric vehicle lithium-ion batteries," *Energies*, vol. 14, no. 22, 2021, Art. no. 7521.
- [5] Z. Deng et al., "General discharge voltage information enabled health evaluation for lithium-ion batteries," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 3, pp. 1295–1306, Jun. 2021.
- [6] Y. Gao et al., "Global parameter sensitivity analysis of electrochemical model for lithium-ion batteries considering aging," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 3, pp. 1283–1294, Jun. 2021.
- [7] Y. Jiang et al., "A review on soft sensors for monitoring, control, and optimization of industrial processes," *IEEE Sensors J.*, vol. 21, no. 11, pp. 12868–12881, Jun. 2021.
- [8] H. Movahedi et al., "Hysteresis compensation and nonlinear observer design for state-of-charge estimation using a nonlinear double-capacitor li-ion battery model," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 1, pp. 594–604, Feb. 2022.
- [9] H. Chen et al., "A single-side neural network-aided canonical correlation analysis with applications to fault diagnosis," *IEEE Trans. Cybern.*, vol. 52, no. 9, pp. 9454–9466, Sep. 2022.
- [10] J. Zhang et al., "Prediction of remaining useful life based on bidirectional gated recurrent unit with temporal self-attention mechanism," *Rel. Eng. System Saf.*, vol. 221, 2022, Art. no. 108297.
- [11] H. Chen et al., "Fault detection for nonlinear dynamic systems with consideration of modeling errors: A data-driven approach," *IEEE Trans. Cybern.*, to be published, doi: 10.1109/TCYB.2022.3163301.
- [12] J. Zhang et al., "An adaptive remaining useful life prediction approach for single battery with unlabeled small sample data and parameter uncertainty," *Rel. Eng. System Saf.*, vol. 222, 2022, Art. no. 108357.
- [13] Y. Qin, J. Zhou, and D. Chen, "Unsupervised health indicator construction by a novel degradation-trend-constrained variational autoencoder and its applications," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 3, pp. 1447–1456, Jun. 2022.
- [14] J. Zhang et al., "Prediction of material removal rate in chemical mechanical polishing via residual convolutional neural network," *Control Eng. Pract.*, vol. 107, 2021, Art. no. 104673.
- [15] H. Chen et al., "Explainable intelligent fault diagnosis for nonlinear dynamic systems: From unsupervised to supervised learning," *TechRxiv*, to be published, doi: 10.36227/techrxiv.19101512.v1.
- [16] Y. Cheng et al., "Auto-encoder quasi-recurrent neural networks for remaining useful life prediction of engineering systems," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 2, pp. 1081–1092, Apr. 2022.
- [17] Z. Lyu, G. Wang, and R. Gao, "Li-ion battery prognostic and health management through an indirect hybrid model," *J. Energy Storage*, vol. 42, 2021, Art. no. 102990.
- [18] B. Gou, Y. Xu, and X. Feng, "State-of-health estimation and remaining-useful-life prediction for lithium-ion battery using a hybrid data-driven method," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 10854–10867, Oct. 2020.
- [19] Q. Miao et al., "Remaining useful life prediction of lithium-ion battery with unscented particle filter technique," *Microelectronics Rel.*, vol. 53, no. 6, pp. 805–810, 2013.
- [20] B. Saha et al., "Prognostics methods for battery health monitoring using a Bayesian framework," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 2, pp. 291–296, Feb. 2009.
- [21] C. Liu, Y. Wang, and Z. Chen, "Degradation model and cycle life prediction for lithium-ion battery used in hybrid energy storage system," *Energy*, vol. 166, pp. 796–806, 2019.
- [22] A. Guha and A. Patra, "State of health estimation of lithium-ion batteries using capacity fade and internal resistance growth models," *IEEE Trans. Transport. Electrification*, vol. 4, no. 1, pp. 135–146, Mar. 2017.
- [23] X. Liu and X. Yang, "Identification of nonlinear state-space systems with skewed measurement noises," *IEEE Trans. Circuits Syst. I: Regular Papers*, to be published, doi: 10.1109/TCSI.2022.3193444.
- [24] Y. Zhou and M. Huang, "Lithium-ion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and arima model," *Microelectronics Rel.*, vol. 65, pp. 265–273, 2016.
- [25] B. E. Olivares et al., "Particle-filtering-based prognosis framework for energy storage devices with a statistical characterization of state-of-health regeneration phenomena," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 2, pp. 364–376, Feb. 2013.
- [26] S.-J. Zhang, R. Kang, and Y.-H. Lin, "Remaining useful life prediction for degradation with recovery phenomenon based on uncertain process," *Rel. Eng. System Saf.*, vol. 208, 2021, Art. no. 107440.
- [27] G. Jin, D. E. Matthews, and Z. Zhou, "A Bayesian framework for on-line degradation assessment and residual life prediction of secondary batteries in spacecraft," *Rel. Eng. Syst. Saf.*, vol. 113, pp. 7–20, 2013.
- [28] K. Park et al., "LSTM-based battery remaining useful life prediction with multi-channel charging profiles," *IEEE Access*, vol. 8, pp. 20786–20798, 2020.
- [29] Q. Ma et al., "Remaining useful life prediction of lithium battery based on capacity regeneration point detection," *Energy*, vol. 234, 2021, Art. no. 121233.
- [30] X. Pang et al., "A lithium-ion battery RUL prediction method considering the capacity regeneration phenomenon," *Energies*, vol. 12, no. 12, 2019, Art. no. 2247.
- [31] Y. Li et al., "Wiener-based remaining useful life prediction of rolling bearings using improved Kalman filtering and adaptive modification," *Measurement*, vol. 182, 2021, Art. no. 109706.
- [32] L. Zhao, Y. Wang, and J. Cheng, "A hybrid method for remaining useful life estimation of lithium-ion battery with regeneration phenomena," *Appl. Sci.*, vol. 9, no. 9, 2019, Art. no. 1890.
- [33] S. Kim et al., "Forecasting state-of-health of lithium-ion batteries using variational long short-term memory with transfer learning," *J. Energy Storage*, vol. 41, 2021, Art. no. 102893.
- [34] N. Zhang et al., "Remaining useful life prediction of lithium batteries based on extended Kalman particle filter," *IEEE Trans. Elect. Electron. Eng.*, vol. 16, no. 2, pp. 206–214, 2021.



- [35] D. Lee and D. Choi, "Analysis of the reliability of a starter-generator using a dynamic Bayesian network," *Rel. Eng. Syst. Saf.*, vol. 195, 2020, Art. no. 106628.
- [36] Y. Huang et al., "A novel robust student's t-Based Kalman filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 3, pp. 1545–1554, Jun. 2017.
- [37] Y. Huang et al., "Robust rauch–tung–striebel smoothing framework for heavy-tailed and/or skew noises," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 415–441, Feb. 2020.
- [38] R. S. M. de Barros, J. I. G. Hidalgo, and D. R. de Lima Cabral, "Wilcoxon rank sum test drift detector," *Neurocomputing*, vol. 275, pp. 1954–1963, 2018.



**Jiushi Zhang** (Student Member, IEEE) received the B.E. degree in automation from Harbin Engineering University, Harbin, China, in 2019, and the M.Sc. degree in control science and engineering from Harbin Institute of Technology, Harbin, China, in 2021, where he is currently working toward the Ph.D. degree in control science and engineering.

His research interests include fault diagnosis and prognosis, predictive maintenance, reliability and safety, prognostics and health management.



**Yuchen Jiang** (Member, IEEE) received the B.E. degree in automation and the Ph.D. degree in control science and engineering from the Harbin Institute of Technology, Harbin, China, in 2016 and 2021, respectively.

He is currently with the School of Astronautics, Harbin Institute of Technology. His research interests include data-driven process monitoring, fault diagnosis and prognosis, industrial cyber-physical systems, and artificial intelligence.



**Xiang Li** (Student Member, IEEE) received the B.E. degree in automation from Northeastern University, Qinhuaogdao, China, in 2019. He is currently working toward the Ph.D. degree in control science and engineering with Harbin Institute of Technology, Harbin, China.

His research interests include medical image analysis and deep learning.



**Hao Luo** (Senior Member, IEEE) received the B.E. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2007, and the M.Sc. and the Ph.D. degrees in electrical engineering and information technology from the University of Duisburg-Essen, Duisburg, Germany, in 2012 and 2016, respectively.

He is currently a Professor with the School of Astronautics, Harbin Institute of Technology, Harbin, China. His research interests include model-based and data-driven fault diagnosis, fault-tolerant systems, and their plug-and-play application on industrial systems.



**Shen Yin** (Senior Member, IEEE) received the M.Sc. degree in control and information system and Ph.D. (Dr.-Ing.) degree in electrical engineering and information technology from the University of Duisburg-Essen, Duisburg, Germany, in 2007 and 2012, respectively.

He is currently a Full Professor (DNV-GL Professor) with the Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway. His research interests include fault diagnosis/prognosis and fault-tolerance strategy, safety, reliability, and security of complicated systems, system and control theory, data-driven and machine learning approaches, and applications in health diagnosis and cyber-physical systems.



**Okyay Kaynak** (Life Fellow, IEEE) received the B.Sc. (first-class Hons.) and the Ph.D. degrees in electronic and electrical engineering from the University of Birmingham, Birmingham, U.K., in 1969 and 1972, respectively.

From 1972 to 1979, he held various positions within the industry. In 1979, he joined Bogazici University, Istanbul, Turkey, where he is currently a Professor Emeritus, holding the UNESCO Chair on Mechatronics. He has held long-term (near to or more than a year) Visiting

Professor/Scholar positions at various institutions in Japan, Germany, USA, Singapore and China. He has authored three books and edited five and authored or coauthored more than 450 papers that have appeared in various journals, books, and conference proceedings. His current research interests include the fields of intelligent control and CPS.

Dr. Kaynak is active in international organizations, has served on many committees of IEEE and was the president of IEEE Industrial Electronics Society during 2002–2003. He was elevated to IEEE fellow status in 2003. In 2016, he received the Chinese Government Friendship Award and Humboldt Research Prize. Most recently, in 2020, he received the International Research Prize of the Turkish Academy of Sciences. He is a member of this Academy.