Supplementary Information

Forecasting battery degradation trajectory under domain shift with domain generalization

Ruifeng Tan[1], Xibin Lu[1,2,3], Minhao Cheng[4], Jia Li[5*],

Jiaqiang Huang[1,2,3*] and Tong-Yi Zhang[6*]

[1]Guangzhou Municipal Key Laboratory of Materials Informatics and Sustainable Energy and Environment Thrust, The Hong Kong University of Science and Technology (Guangzhou), Nansha, Guangzhou, 511400, Guangdong, P.R. China.

[2]Academy of Interdisciplinary Studies, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, 999077, Hong Kong, P.R. China.

[3]HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, Shenzhen, 518045, Guangdong, P.R. China.

[4]Department of Computer Science & Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, 999077, Hong Kong, P.R. China.

[5]Guangzhou Municipal Key Laboratory of Materials Informatics and Data Science and Analytics Thrust, The Hong Kong University of Science and Technology (Guangzhou), Nansha, Guangzhou, 511400, Guangdong, P.R. China.

[6]Guangzhou Municipal Key Laboratory of Materials Informatics, Advanced Materials Thrust, and Sustainable Energy and Environment Thrust, The Hong Kong University of Science and Technology (Guangzhou), Nansha, Guangzhou, 511400, Guangdong, P.R. China.

*Corresponding authors.

E-mails: jialee@ust.hk; seejhuang@ust.hk; mezhangt@hkust-gz.edu.cn

Supplementary Notes

Supplementary Note 1. Selection of L_h and L_f

Application-wise, L_h should be as small as possible, whereas the L_f should be as large as possible. However, the performance of model usually deteriorates with a smaller L_h and a larger L_f. Consequently, we decided to take a L_h that is ≤5% of the median cycle lives and a L_f that is >50% of the median cycle lives.

After the data cleaning, the median cycle lives in NCA and LFP datasets are 200 and 729, respectively. According to the above principle, we chose 10 and 20 as the L_h for the NCA and LFP datasets, respectively. Meanwhile, 150 and 500 were selected as the L_f for the NCA and LFP datasets, respectively.

Supplementary Note 2. Correlation analysis between remaining useful life and forecasted states

We calculated the Pearson correlation coefficients between the remaining useful life (RUL) and the forecasted states (capacity and energy). The Pearson correlation coefficients between the RUL and energy range from 0.91/0.98 to 0.99/0.99, and the counterpart between the RUL and capacity range from 0.89/0.98 to 0.99/0.99 for the LFP/NCA dataset.

Supplementary Note 3. Data preprocessing

It is observed that both NCA[1] and LFP[2] datasets contain some abnormal data. Most of them are caused by missing data and experimental errors. Here, we illustrate how we process the raw data to get the data used in this work.

For the NCA dataset, we remove the cells that haven't degraded to 80% of nominal capacity. Additionally, we remove the data after the first life (80% of nominal capacity). After that, the first-life data of each cell is processed according to the following steps:

1. We abandon the cycle that has a large period of missing time records. We guess some unknown tests were conducted during the missing data. The next cycle that is affected by the unknown test is removed.
2. We abandon the cycle that has obviously abnormal energy efficiency(ee), such as ee<=0.8 or ee>=1.5.
3. We abandon the cycle that has abnormal discharging/charging capacity(d/cc), such as d/cc<=2300 mAh. Moreover, we remove the cycle that has a sudden

charging capacity drop compared with the previous cycle, such as a drop larger than 100 mAh.

4. After the above operations, we make the cycle number continuous and save the processed data.

For the LFP dataset, we first process it using the official processing Python code, and then process each cell according to the following procedures:

1. We remove the data after the first life (80% of nominal capacity).
2. We remove the cycle that has a sudden change of current during discharging when the current should be designed to be constant.
3. We remove the cycle that has multiple charging and discharging.
4. We remove the cycle that has abnormal energy efficiency(ee), such as ee=<0.8 and ee>=1.5.

All the statistics in this work is conducted based on the processed dataset.

Supplementary Note 4. Implementation details of baselines

The Informer[3] and One-shot LSTM[4] are implemented based on their official source codes. For the hyperparameters, we set the learning rate to 0.005 and 0.0075 for the NCA and LFP datasets, respectively, the same as the meta learning rate of MAGNet. Additionally, the searching range of the hyperparameters is the same for all models, including the MAGNet. Specifically, we tune the number of encoder layer and decoder layer from [1,2], the neuron number of the hidden linear layer from [2,4], the weight decay in the AdamW[5] optimizer from [0,0.0001], and keep the rest the same as MAGNet. As for the BiLSTM, we implement it using a two-layer bidirectional LSTM whose hidden size is the same as the embedding dimension of features. The rest of the hyperparameters of BiLSTM are the same as MAGNet. The best hyperparameters are chosen according to the performances on validation sets.

Supplementary Note 5. Details of the calculation of the evaluation metrics

Three evaluation metrics are adopted to evaluate the model performance: α-accuracy[6], MAE, and R^2. Among them, α-accuracy is originally proposed to assess the model uncertainty from an engineering perspective[7], and it has been introduced to evaluate the model performance for battery SOH estimation by Roman et al.[6] It represents the percentage of predictions that lie in an accuracy zone defined by α (α can be adjusted according to the intended applications), where α is the threshold of the relative percentage error deviating from the true values. For the whole testing set, it is calculated as follows:

$$\alpha - \text{accurcay} = \frac{1}{Q}\sum_{i=1}^{Q}\frac{1}{N_i}\sum_{j=1}^{N_i}\frac{1}{N_j^i}\sum_{k=1}^{N_j^i}\mathcal{A}_k^{(j)}, \quad (1)$$

where N_i is the number of cells in i^{th} domain, N_j^i means the number of samples in the j^{th} cell of i^{th} domain, and $\mathcal{A}_k^{(j)}$ is the α-accuracy of the k^{th} sample of the j^{th} cell. To further illustrate N_j^i and $\mathcal{A}_k^{(j)}$, let's assume we have a cell that has 550 cycles, $L_h = 20$, and $L_f = 500$. Then, this cell has $N_j = 550 - L_h = 530$ samples. As for $\mathcal{A}_k^{(j)}$, for instance, $\mathcal{A}_1^{(j)}$ is the α-accuracy between the predicted and true values of states from 21st cycle to 520th cycle, and $\mathcal{A}_{530}^{(j)}$ means the α-accuracy between the predicted and true values of states of the 550th cycle. We include the samples that don't have L_f cycles in the evaluation so that the model performance in forecasting the degradation trajectories around the end of life is well reflected. This can avoid evaluation bias on the model performance in forecasting the degradation trajectories around the beginning of life.

The other two metrics are commonly used: MAE and R^2. For total Q domains in the testing set, the MAE is computed as follows:

$$\text{MAE} = \frac{1}{Q}\sum_{i=1}^{Q}\frac{1}{N_i}\sum_{j=1}^{N_i}\frac{1}{N_j^i}\sum_{k=1}^{N_j^i}\mathcal{M}_k^{(j)}, \quad (2)$$

where $\mathcal{M}_k^{(j)}$ denotes the mean absolute error of the k^{th} sample of the j^{th} cell. As for R^2, its computation is:

$$R^2 = \frac{1}{Q}\sum_{i=1}^{Q}\frac{1}{N_i}\sum_{j=1}^{N_i}R_j, \quad (3)$$

where R_j is the R^2 computed based on the predictions and truth of the samples of the j^{th} cell in the i^{th} domain. It should be noted that Equation (3) is slightly different from Equation (2) and Equation (1) since it is not reliable to compute vanilla R^2 when the k^{th} sample only has few cycles. Therefore, for each cycle in a cell, we average the predicted values of a state on it so that we get the averaged predictions of all cycles. After that, for the j^{th} cell, we can use the averaged predictions and truth of each cycle to compute R_j as follows:

$$R_j = 1 - \frac{SS_{res}}{SS_{tot}}, \quad (4)$$

where SS_{res} is the sum of sum of squares of residuals, and SS_{tot} is the total sum of squares.

Supplementary Note 6. The preparation of half cells

First, we disassembled the fresh battery (full cell) sample in an argon-filled glove box to collect the cathode and anode films respectively. Since the active materials are coated on both sides of the films of commercial batteries, the single-side coated films were prepared after carefully wiping out the other side. After cleaning in DMC, the single-side coated cathode and anode films were punched into disks with a diameter of 12 mm. The obtained NCA (cathode) and Si/C anode disks are separately fabricated to coin cells with the Li metal disk as counter electrodes. The separator is about 25 μm thick (Celgard), and the electrolyte recipe used in the coin cells is 1 M $LiPF_6$ in EC: DMC = 1: 2). The above entire process has been carried out in an argon-filled glove box to avoid the exposure to moisture and air.

The prepared coin cells were tested in galvanostatic charge and discharge by high-accurate cycle channels with the accuracy of 0.01% FS. The coin cells of NCA and Si/C were tested at 0.1 C within the voltage range of 2.65~4.25 V and 0.010~1.000 V, respectively. To acquire the differential voltage analysis--dU/dQ of equilibrium potentials, the equilibrium potentials of the cathode ($U_{cathode}$) or anode (U_{anode}) were firstly calculated in the same way in Methods. The theoretical U of the full cell can then be obtained based on the equilibrium potentials of anode (U_{anode}) and equilibrium potentials of cathode ($U_{cathode}$):

$$U_{full\ cell} = U_{cathode} - U_{anode}, \quad (5)$$

After that, we can get the simulated U vs. Q under five different aging modes[1]. Then, the dQ/dU and dU/dQ can be obtained to estimate the degradation modes (see supplementary Fig. 8-19). It is worth noting that slight deviations may exist in the starting/end point of U of the cathode and anode in Supplementary Fig. 8-19, due to the difference between half-coin cells and full cells. However, these deviations will not affect the variation trend of peak positions compared to pristine battery. Since the original dU/dQ and dQ/dU are noisy, we used the voltage window method to get the smoothed curves. Specifically, we got the voltage and capacity value every 0.005V using linear interpolations from original data, and then the dU/dQ and dQ/dU can be computed as:

$$dU/dQ = \triangle U/\triangle Q, \quad (6)$$
$$dQ/dU = \triangle Q/\triangle U. \quad (7)$$

Supplementary Note 7. Estimation method for degradation modes in the NCA dataset

In this note, we introduce the way to quantify the degradation degree of LLI and

$LAM_{PE,De}$. For LLI, we simulated the U v.s. Q plots under different degradation degrees of different degradation modes based on the work of Birkl et al[8]. It is observed from Supplementary Fig. 20 that LLI clearly affects the capacity between 3.4 and 3.6V ($Cap_{3.4V-3.6V}$). At the same time, the effects of $LAM_{PE,De}$ and $LAM_{NE,De}$ on $Cap_{3.4V-3.6V}$ are negligible. Thus, we assume that $Cap_{3.4V-3.6V}$ is only affected by LLI. $Cap_{3.4V-3.6V}$ versus the degradation degree of LLI is shown in Supplementary Fig. 20. Because the capacity of the experimental full cell in the NCA dataset is different from the simulated full cell, we normalize the $Cap_{3.4V-3.6V}$ by dividing each value with $Cap_{3.4V-3.6V}$ at the pristine state (normalized $Cap_{3.4V-3.6V}$ versus the degradation degree of LLI is shown in Supplementary Fig. 21). Then, for each cycle in an experimental full cell, we can get the degradation degree of LLI in it based on the linear interpolation of its normalized $Cap_{3.4V-3.6V}$.

For $LAM_{PE,De}$, we observe from the Supplementary Fig. 16 that the distance between the distance between the M and H2 peaks ($Dist_{M-H2}$) will be smaller when $LAM_{PE,De}$ happens. At the same time, according to Supplementary Fig. 15, LLI generally doesn't affect $Dist_{M-H2}$ since LLI only causes the shift of anode U versus Q curve. Therefore, we use the change ratio of $Dist_{M-H2}$ to measure the degradation degree of $LAM_{PE,De}$ for the experimental cells in the NCA dataset:

$$\text{Degradation degree of } LAM_{PE,De} = 1 - \frac{Dist_{M-H2}^{N}}{Dist_{M-H2}^{1}}, \tag{8}$$

where $Dist_{M-H2}^{N}$ is the $Dist_{M-H2}$ at the N^{th} cycle, and $Dist_{M-H2}^{1}$ is the $Dist_{M-H2}$ at the first cycle. It should be noted that the dU/dQ curves are smoothed so that $Dist_{M-H2}$ can be measured. For an experimental full cell in the NCA dataset, we use the *smoothdata* (https://ww2.mathworks.cn/help/matlab/ref/smoothdata.html) function in MATLAB to smooth the raw $\triangle U/\triangle Q$ curves directly computed from the data. The selected method is Gaussian smoothing, and the window size is 15. After that, we can get the cubic interpolation of dU/dQ w.r.t Q. Thus, we can get the dU/dQ value at the capacity range with a unit of 1mAh so that the smoothed dU/dQ plot is obtained. We also observe that the position of M peak or H2 peak sometimes becomes unclear for an operating condition. Therefore, we might use the distance between the M/H2 peak and the M-H2 valley (the valley between M and H2 peaks in the dU/dQ plot) to replace the $Dist_{M-H2}$ term in the Equation (8) to compute the degradation degree of $LAM_{PE,De}$ in its whole life. The detailed choice of the distance for each condition is indicated by the yellow arrows in Fig. 5e. Take the 45℃-0.5C as an example, we

observe that the M peak might be hard to identify since it merges with anode peak around the end of life. Then, we will use the distance between H2 peak and the M-H2 valley to quantify its $LAM_{PE,De}$ throughout its life. It should be noted both M and H2 peaks are hard to be detected for 25-1 condition. In this case, we find that the capacity between 3.75 and 3.85V ($Cap_{3.75V-3.85V}$) is mainly affected by the $LAM_{PE,De}$ (Supplementary Fig. 24). The decreased $Cap_{3.75V-3.85V}$ in an experimental NCA full cell well corresponds to the dQ/dU cathode peak loss between 3.75V and 3.85V in the 25-1 condition, thus can be utilized to quantify the degradation degree of $LAM_{PE,De}$ for it. Specifically, we normalize the $Cap_{3.75V-3.85V}$ by dividing each value with $Cap_{3.75V-3.85V}$ at the pristine state, and show its relationship to the degradation degree of $LAM_{PE,De}$ in Supplementary Fig. 25. Then, for each cycle in an experimental full cell, we can get the degradation degree of $LAM_{PE,De}$ in it based on the linear interpolation of its normalized $Cap_{3.75V-3.85V}$.

Supplementary Note 8. Computation of MAGNet superiority

To study if the space of the encoder embeddings of MAGNet is more similar to the mechanism space compared to the counterpart of Informer, we first measure the distance between different operating conditions for each space. Specifically, we standard normalize the feature value for each sample and then measure the Wasserstein distance between different operating conditions to form a distance matrix A_{ij}, where the value in the i^{th} row and j^{th} column is the Wasserstein distance between condition i and condition j. After that, we get the absolute value of the difference between the distance matrix of the mechanism space $A_{ij}^{Mechanism}$ and distance matrix of the MAGNet/Informer space $A_{ij}^{MAGNet}/A_{ij}^{Informer}$:

$$D_{ij}^{MAGNet} = abs\left(A_{ij}^{MAGNet} - A_{ij}^{Mechanism}\right), \tag{9}$$

$$D_{ij}^{Informer} = abs\left(A_{ij}^{Informer} - A_{ij}^{Mechanism}\right), \tag{10}$$

where D_{ij}^{MAGNet} is the distance between the MAGNet space and mechanism space w.r.t. condition i and condition j. $D_{ij}^{Informer}$ is the distance between the Informer space and mechanism space w.r.t. condition i and condition j. $abs(\cdot)$ is a function that sets the value as the absolute value. Then, we can get the superiority of MAGNet over Informer as follows:

$$S_{ij} = D_{ij}^{Informer} - D_{ij}^{MAGNet} \tag{11}$$

where S_{ij} is the superiority of MAGNet over Informer w.r.t condition i and condition

j. Because S_{ij} is symmetric, we only use lower triangle of it in Fig. 5g. The Wasserstein distance is computed using the wasserstein package in Python (https://pkomiske.github.io/Wasserstein/docs/emd/). The overall distance between two spaces is computed as:

$$O^{MAGNet} = \left|\left|D^{MAGNet}\right|\right|_F \tag{12}$$

$$O^{Informer} = ||D^{Informer}||_F \tag{13}$$

where O^{MAGNet} is the overall distance between the MAGNet space and mechanism space, $O^{Informer}$ is the overall distance between the Informer space and mechanism space, and $||\cdot||_F$ is the Frobenius norm.

Supplementary Tables

Supplementary Table 1. Details of the NCA dataset split.

Train/Val/Test	Conditions	The number of cells	Averaged cycle life	Domain
Train	25℃-0.25C	3	218.3	Source
	25℃-0.5C	10	141.2	
	35℃-0.5C	1	367.0	
	45℃-0.5C	14	318.4	
	25℃-1C	5	22.6	Target
Val	25℃-0.25C	1	194	Source
	25℃-0.5C	3	152.7	
	35℃-0.5C	0	0	
	45℃-0.5C	6	376.8	
	25℃-1C	2	20	Target
Test	25℃-0.25C	1	200	Source
	25℃-0.5C	4	148	
	35℃-0.5C	1	403	
	45℃-0.5C	5	338	
	25℃-1C	2	22.5	Target

Supplementary Table 2. Details of the LFP dataset split.

Row index	Conditions	The number of cells	Averaged cycle life	Domain	Train/Val/Test
1	8C(35%)-3.6C	2	488.5	Source	Train
2	8C(25%)-3.6C	2	618.5	Source	Train
3	8C(15%)-3.6C	2	945.5	Source	Train
4	7C(40%)-3C	2	635.5	Source	Train
5	7C(40%)-3.6C	2	582	Source	Train
6	7C(30%)-3.6C	2	690	Source	Train
7	6C(50%)-3.6C	2	790.5	Source	Train
8	6C(40%)-3C	2	925	Source	Train
9	6C(40%)-3.6C	2	854	Source	Train
10	6C(4%)-4.75C	1	484	Source	Train
11	6C(30%)-3.6C	1	1011	Source	Train
12	5C(67%)-4C	6	1113.5	Source	Train
13	5.9C(60%)-3.1C	2	864.5	Source	Train
14	5.9C(15%)-4.6C	2	865	Source	Train
15	5.6C(5%)-4.75C	1	497	Source	Train
16	5.6C(38%)-4.25C	1	475	Source	Train

17	5.6C(36%)-4.3C	8	928.875	Source	Train
18	5.4C(70%)-3C	2	711.5	Source	Train
19	5.4C(60%)-3C	2	796.5	Source	Train
20	5.4C(50%)-3C	1	785	Source	Train
21	5.4C(40%)-3.6C	1	1051	Source	Train
22	5.3C(54%)-4C	8	1072.125	Source	Train
23	5.2C(58%)-4C	1	517	Source	Train
24	5.2C(50%)-4.25C	1	516	Source	Train
25	5.2C(37%)-4.5C	1	489	Source	Train
26	5.2C(10%)-4.75C	1	479	Source	Train
27	4C(80%)-4C	2	1567	Source	Train
28	4C(40%)-6C	1	480	Source	Train
29	4C(4%)-4.85C	1	483	Source	Train
30	4.9C(69%)-4.25C	1	495	Source	Train
31	4.9C(61%)-4.5C	1	507	Source	Train
32	4.65C(44%)-5C	1	510	Source	Train
33	4.65C(19%)-4.85C	1	486	Source	Train
34	4.4C(80%)-4.4C	1	1070	Source	Train
35	4.4C(47%)-5.5C	1	485	Source	Train
36	4.4C(24%)-5C	1	491	Source	Train
37	3.7C(31%)-5.9C	3	657.66	Source	Train
38	3.6C(9%)-5C	1	558	Source	Train
39	3.6C(80%)-3.6C	3	2077.66	Source	Train
40	3.6C(30%)-6C	1	508	Source	Train
41	5.4C(60%)-3.6C	2	857	Source	Val
42	4.4C(8%)-4.85C	1	499	Source	Val
43	5.2C(66%)-3.5C	1	495	Source	Val

44	3.6C(2%)-4.85C	1	477	Source	Val
45	5.6C(19%)-4.6C	7	961.14	Source	Val
46	4.8C(80%)-4.8C	9	1029.11	Source	Val
47	4C(13%)-5C	1	474	Source	Val
48	5.4C(80%)-5.4C	2	543.5	Source	Val
49	4.65C(69%)-6C	1	524	Source	Val
50	6C(60%)-3C	3	730	Source	Val
51	6C(50%)-3C	2	886.5	Source	Val
52	5.6C(25%)-4.5C	1	532	Source	Val
53	6C(52%)-3.5C	1	427	Target	Test
54	6C(40%)-4C	1	454	Target	Test
55	6C(31%)-4.25C	1	459	Target	Test
56	6C(20%)-4.5C	1	463	Target	Test
57	5.6C(65%)-3C	1	426	Target	Test
58	5.6C(58%)-3.5C	1	456	Target	Test
59	5.6C(47%)-4C	1	462	Target	Test
60	5.2C(71%)-3C	1	461	Target	Test
61	4C(31%)-5C	1	455	Target	Test
62	4.9C(27%)-4.75C	1	464	Target	Test
63	4.4C(55%)-6C	1	458	Target	Test
64	3.6C(22%)-5.5C	1	441	Target	Test
65	2C(7%)-5.5C	1	332	Target	Test
66	2C(2%)-5C	1	435	Target	Test
67	2C(10%)-6C	1	146	Target	Test
68	1C(4%)-6C	1	297	Target	Test

Supplementary Table 3. The model performance of MAGNet and compared baselines on the NCA dataset. The best results are shown in the bold text, and the second-best results are underlined. It shows that MAGNet consistently outperforms the baselines. The α is 2 here.

	Capacity			**Energy**		
	α-accuracy(%)	R^2	MAE(Ah)	α-accuracy(%)	R^2	MAE(Wh)
MAGNet(Ours)	**96.51**	**0.976**	**0.0165**	**93.81**	**0.969**	**0.0637**
Informer	<u>90.69</u>	0.848	<u>0.0252</u>	<u>86.07</u>	0.845	<u>0.0980</u>
One-shot LSTM	86.34	0.848	0.0326	78.45	0.810	0.1282
BiLSTM	85.45	<u>0.881</u>	0.0302	78.97	<u>0.856</u>	0.1185

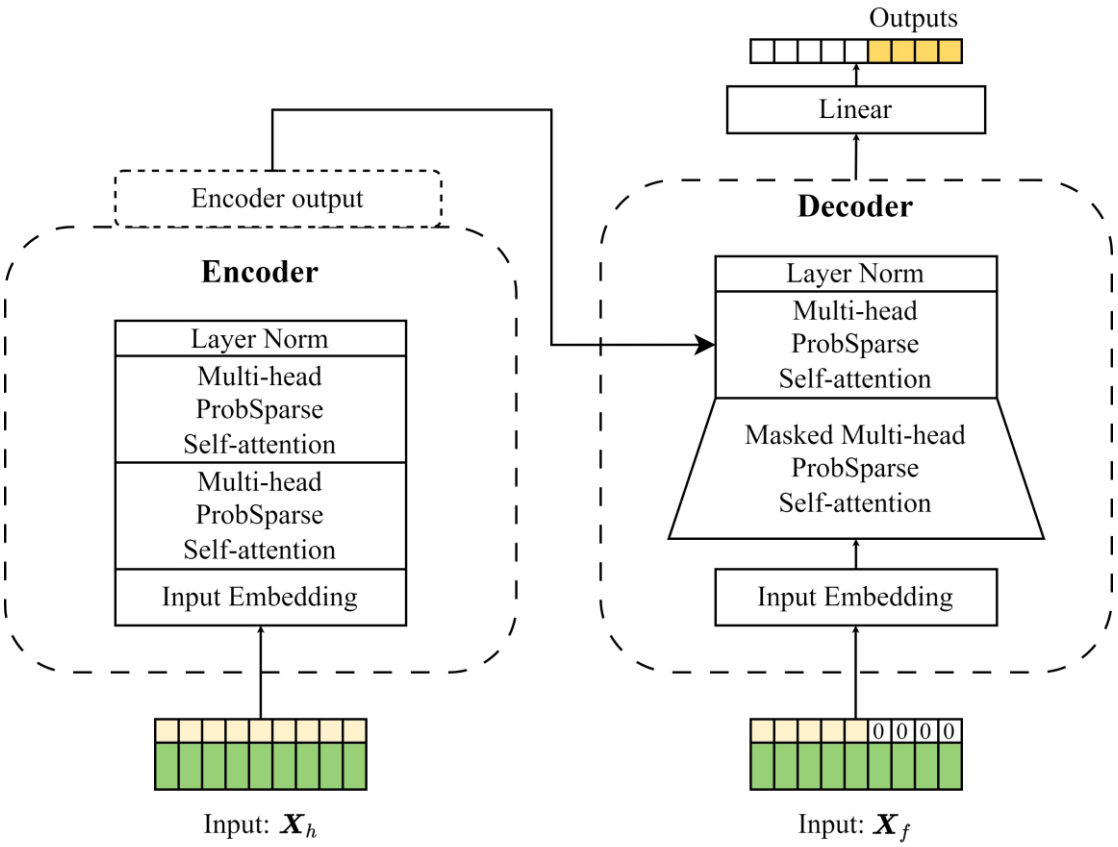
Supplementary Table 4. The model performance of MAGNet and compared baselines on the LFP dataset. The best results are shown in the bold text, and the second-best results are underlined. It shows that MAGNet consistently outperforms the baselines. The α is 4 here.

	Capacity			**Energy**		
	α-accuracy(%)	R^2	MAE(Ah)	α-accuracy(%)	R^2	MAE(Wh)
MAGNet(Ours)	**88.11**	**0.891**	**0.0154**	**84.08**	**0.902**	**0.0512**
Informer	<u>78.44</u>	0.762	<u>0.0227</u>	<u>71.68</u>	<u>0.787</u>	<u>0.0775</u>
One-shot LSTM	75.18	<u>0.771</u>	0.0269	61.67	0.653	0.1153
BiLSTM	72.57	0.649	0.0308	57.01	0.511	0.1300

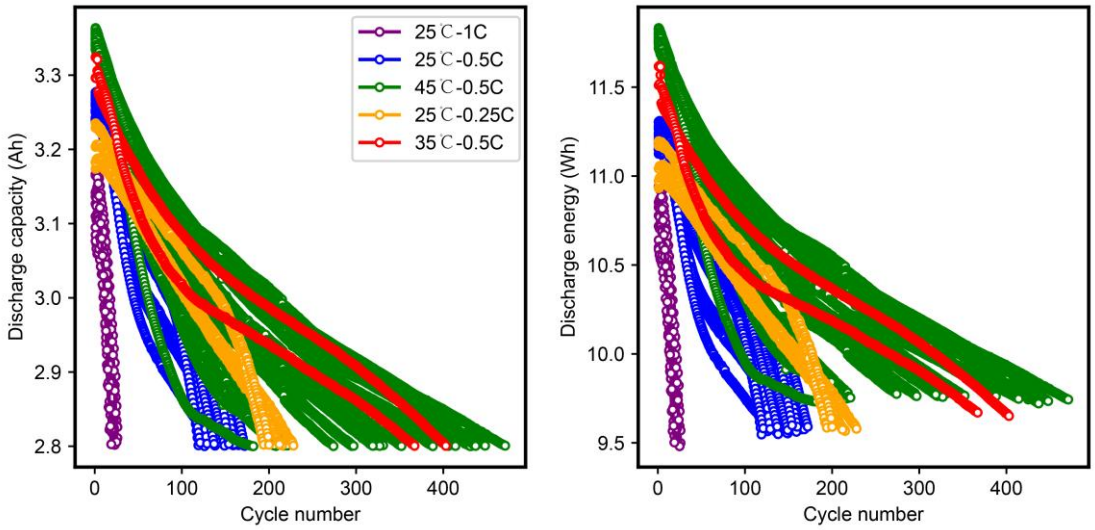
Supplementary Table 5. The model core configuration for NCA and LFP datasets.

	NCA	**LFP**
The number of parameters	2017	5083
The embedding dimension of the features	8	12
The number of attention layers in the Encoder	1	2
The number of attention layers in the Decoder	2	2
The activation function	GELU	GELU
The attention factor for the attention layer in the Encoder	4	5
The attention factor for the attention layer in Decoder	1	1
The neuron number of the hidden linear layers	2	4
β in the loss function	0.25	2
γ in the loss function	0.25	0.2
Batch size	32	128
Learning rate λ	10−6	10−6
Meta learning rate	0.005	0.0075

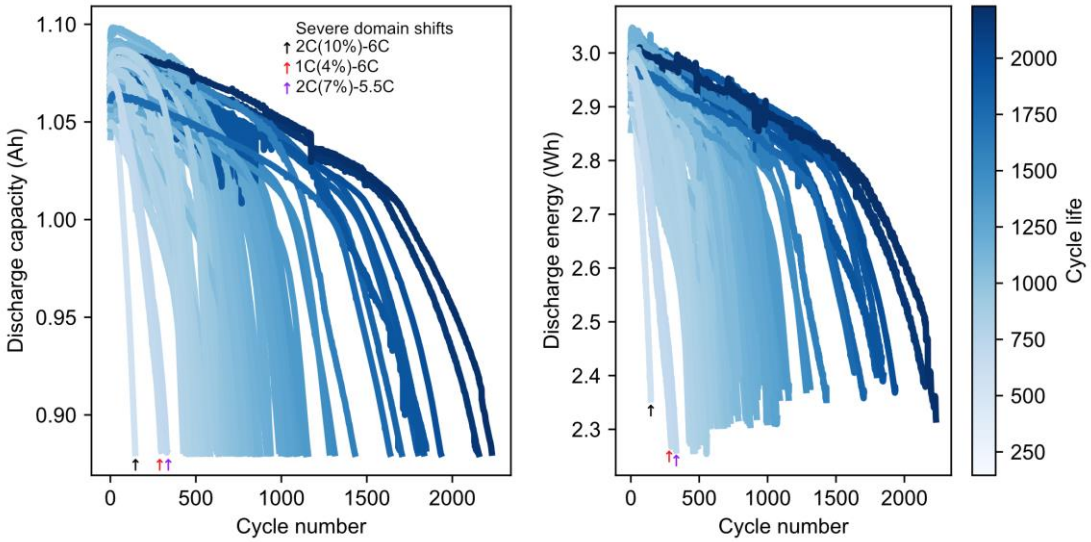
Supplementary Figures



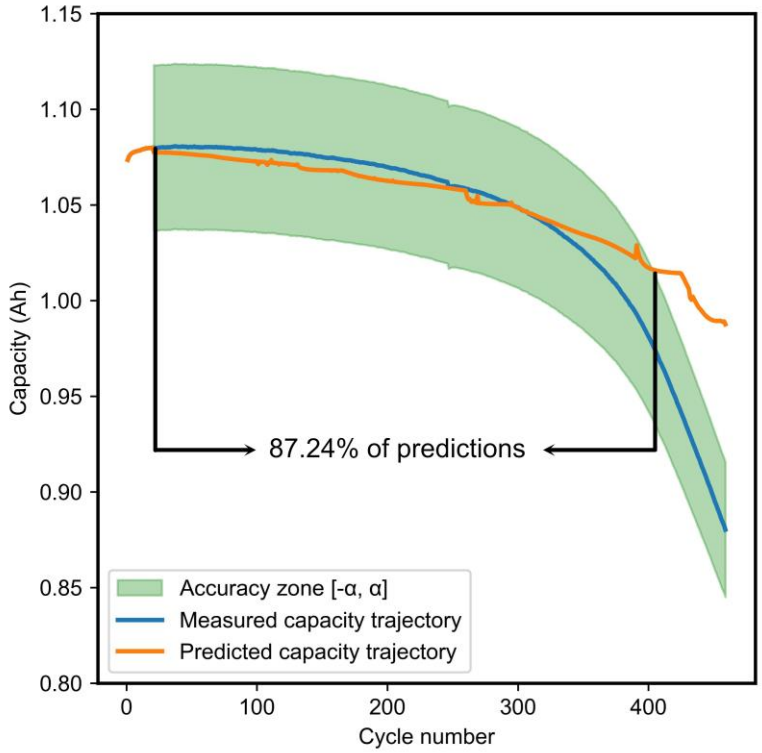
Supplementary Figure 1. The details of the Informer encoder and decoder. The same encoder and decoder are adopted in MAGNet.



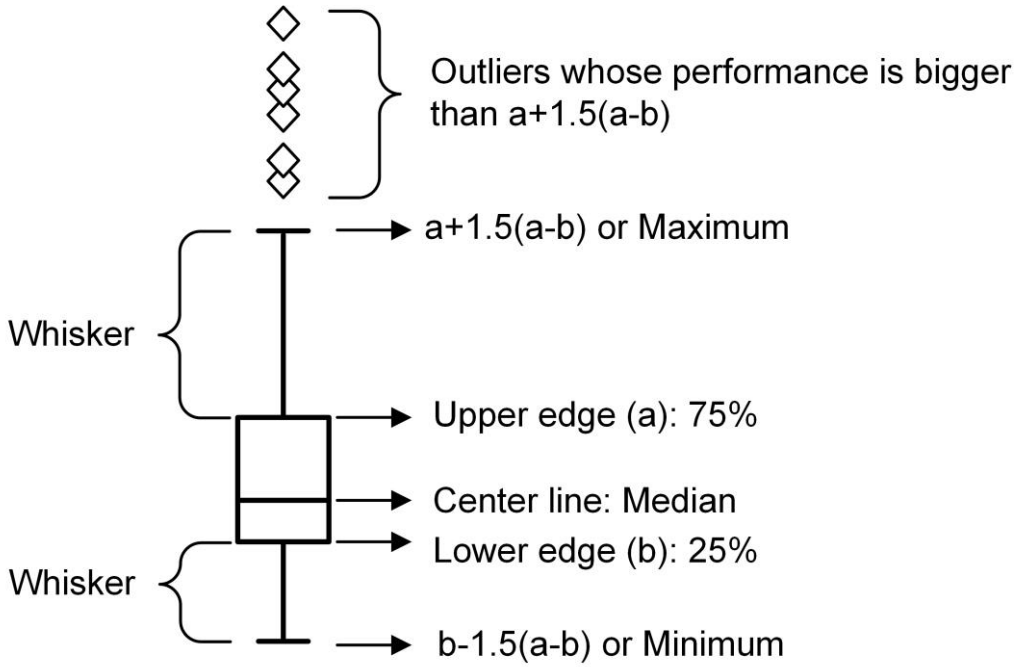
Supplementary Figure 2. The capacity and energy trajectories of all the cells in the processed NCA dataset. For AA°C-BBC, AA°C means the temperature, BBC is the charging rate. The discharging rate is 1C for all conditions. It shows that the trajectories of different conditions are different, and even the trajectories in the same condition can vary from each other. The capacity and energy trajectories are similar but different.



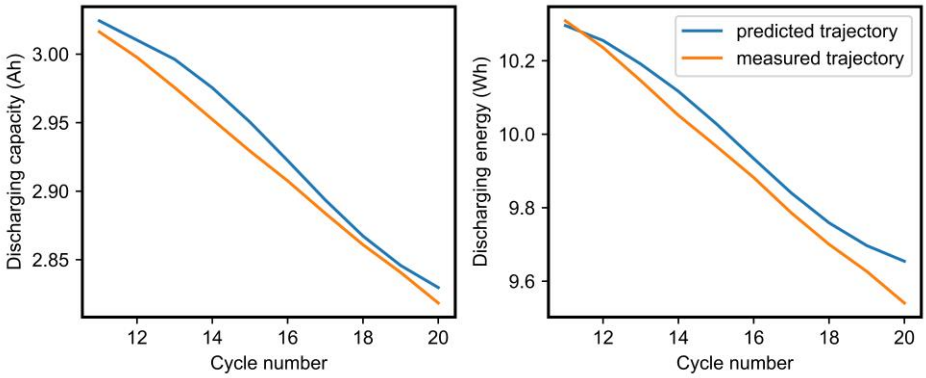
Supplementary Figure 3. The capacity and energy trajectories of all the cells in the processed LFP dataset. The color lightness indicates the cycle life. It shows that the trajectories can vary greatly from each other.



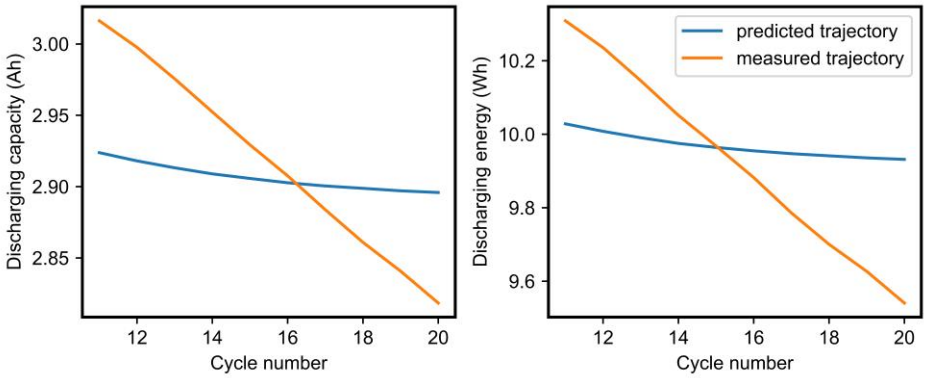
Supplementary Figure 4. Illustration α-accuracy. The α-accuracy is the percentage of predictions that lie within the green accuracy zone that is defined by α. This example is produced by MAGNet on a testing cell collected from 6C(31%)-4.25C condition. The α is 4. MAGNet receives the first 20 cycles as the input and predicts the 500 cycles in the future. Since the cycle life of this cell is only 459, we only use the predictions between 21 and 459 to calculate the α-accuracy, which is 87.24% here.



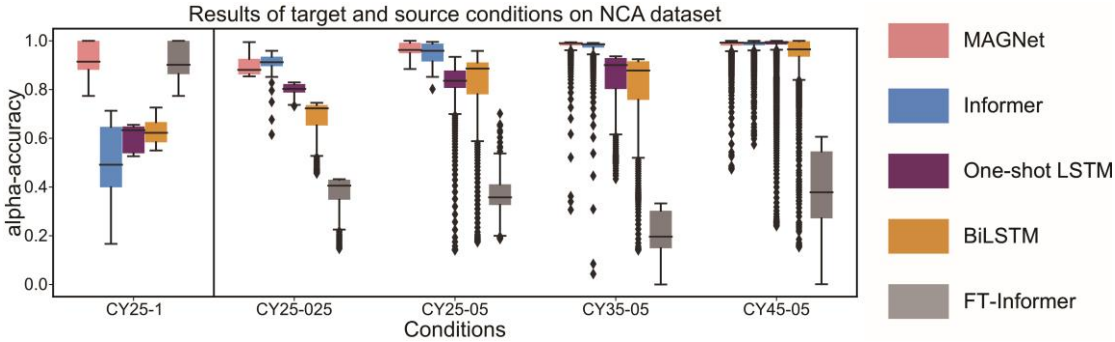
Supplementary Figure 5. Definition of the Boxplot. The length of the whiskers can extend up to 1.5 times the interquartile range, represented by 1.5(a-b) if there exist samples with performance metrics exceeding a+1.5(a-b) or falling below b-1.5(a-b). In the absence of such outliers, the whiskers demarcate the maximum and minimum performance levels of the samples.



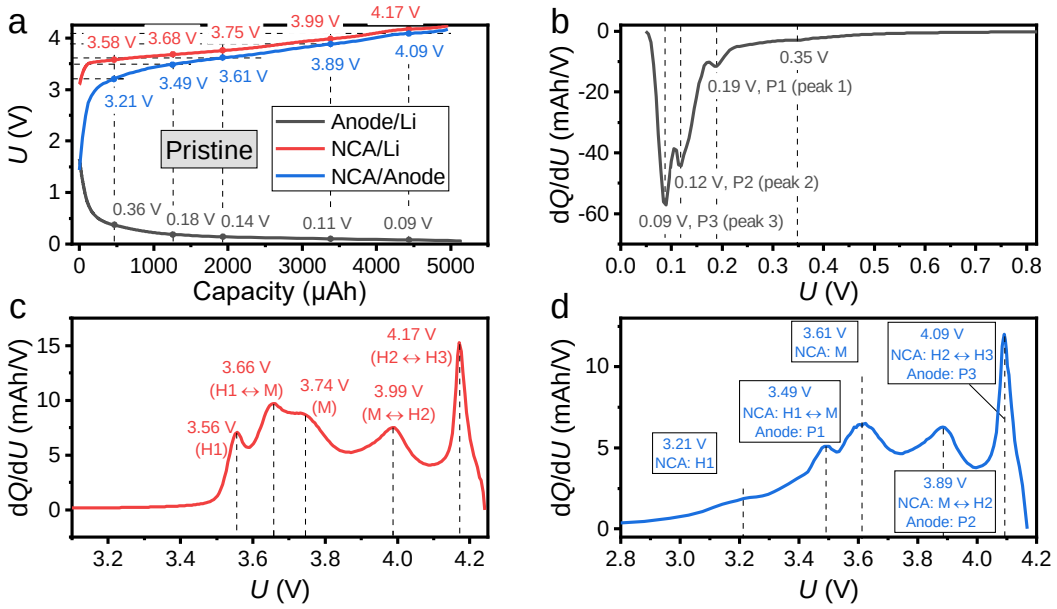
Supplementary Figure 6. The prediction of MAGNet on the number 5 cell from 25℃-1C condition of the NCA dataset based on the first 10 cycles. It shows that this cell degrades to 80% nominal capacity within only 20 cycles. The blue and orange lines are the predicted trajectory and experimentally measured trajectory, respectively. Although most of the training cells have cycle lives much longer than 20, MAGNet can still forecast the fast failure of this cell accurately.



Supplementary Figure 7. The prediction of Informer on the number 5 cell from 25℃-1C condition of NCA dataset based on the first 10 cycles. It shows that the Informer is greatly affected by the relatively long-lived cells that occupy the most proportion of the training data, thus tending to forecast that the capacity of this cell won't quickly degrade.



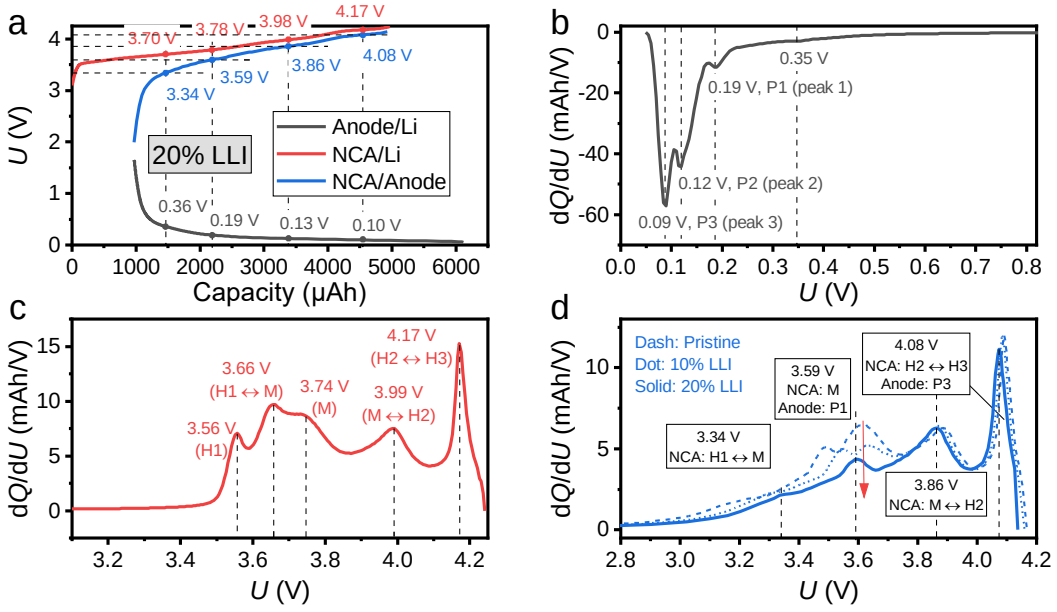
Supplementary Figure 8. The detailed results of all the baselines and MAGNet on each condition of the NCA dataset.



277

278 **Supplementary Figure 9. d*Q*/d*U* plot at the pristine state. a**, Capacity vs. equilibrium potential
279 (*U*) curves of anode/Li (grey), NCA/Li (red), and NCA/anode (blue). **b-d**, The d*Q*/d*U* plots of
280 anode/Li (grey, **b**), NCA/Li (red, **c**), and NCA/anode (blue, **d**). The blue dots in **a** correspond to the
281 peak positions in **d**, which help to identify the redox pairs in anode (**b**) and NCA (**c**) corresponding
282 to the full cell (**d**).

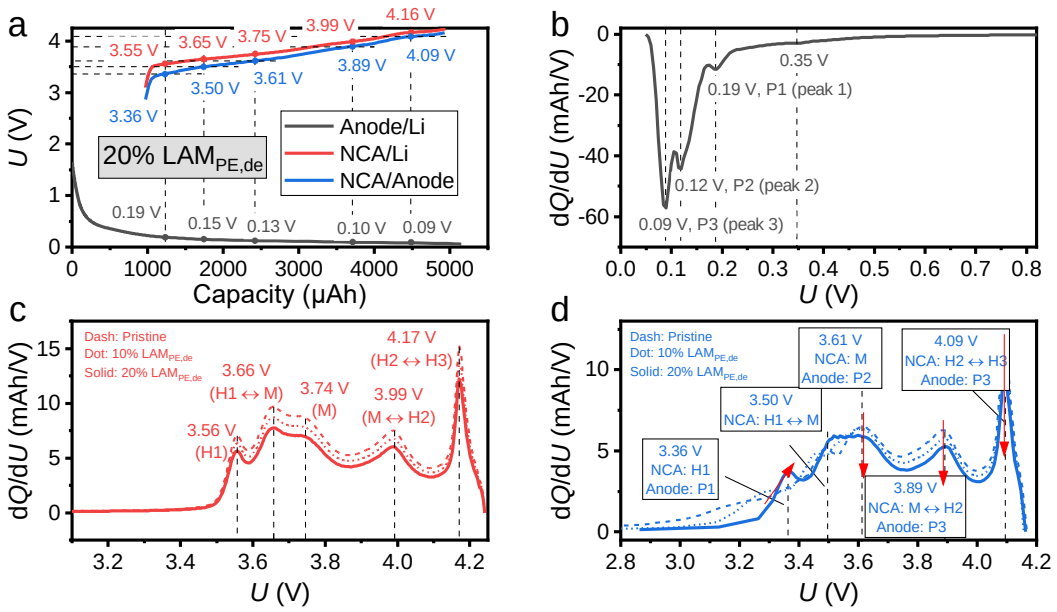
283



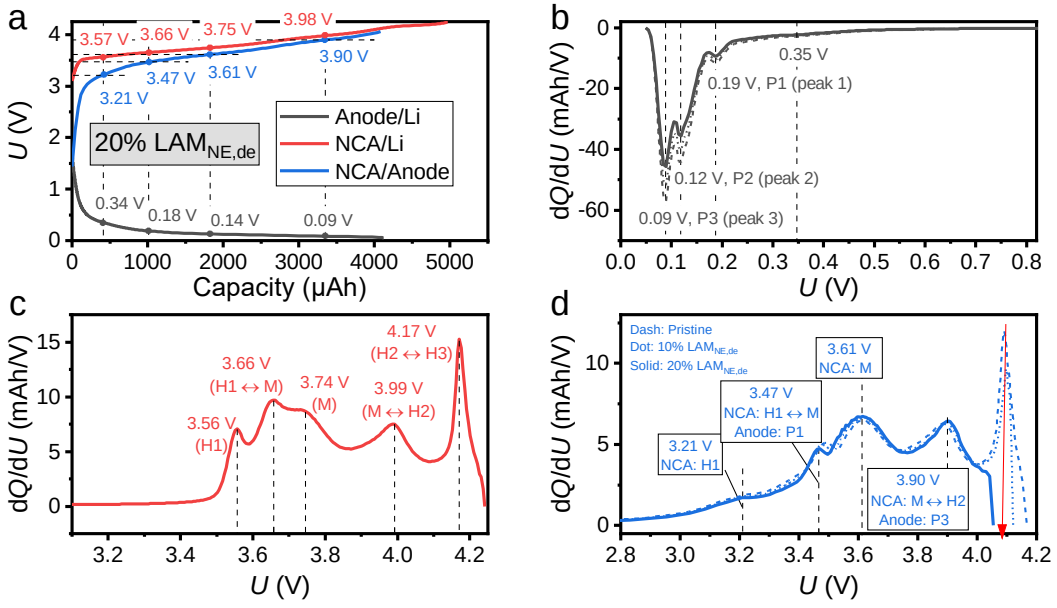
284

285 **Supplementary Figure 10. d*Q*/d*U* plot with 20% loss of lithium inventory (LLI). a**, Capacity
286 vs. equilibrium potential (*U*) curves of anode/Li (grey), NCA/Li (red), and NCA/anode (blue). **b-d**,
287 The d*Q*/d*U* plots of anode/Li (grey, **b**), NCA/Li (red, **c**), and NCA/anode (blue, **d**). The blue dots in
288 **a** correspond to the peak positions in **d**, which help to identify the redox pairs in anode (**b**) and NCA
289 (**c**) corresponds to the full cell (**d**). The dash, dot, and solid lines in **d** indicate the pristine, 10% LLI,
290 and 20% LLI, respectively.

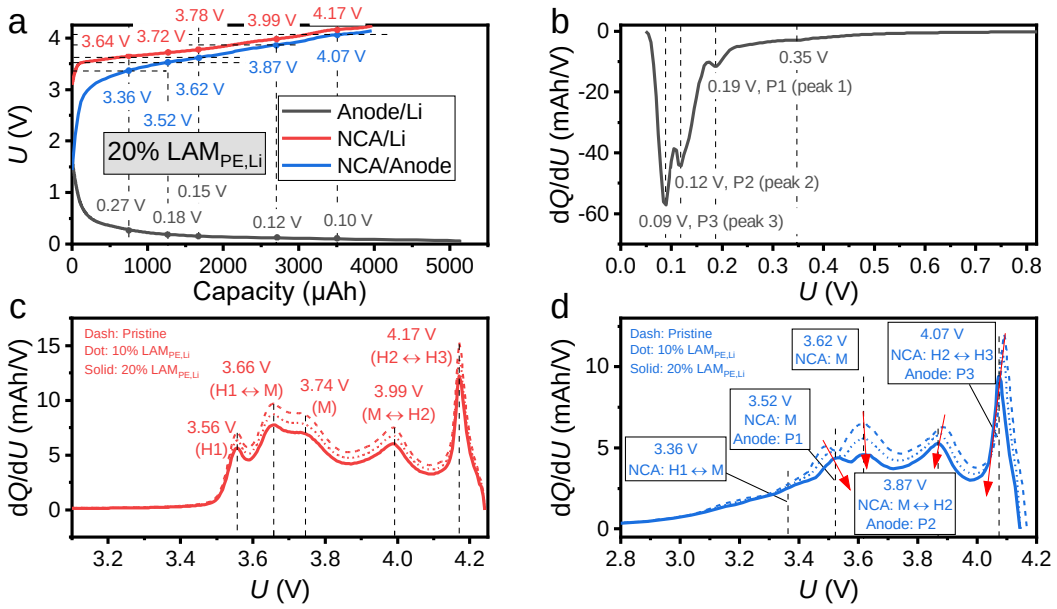
291



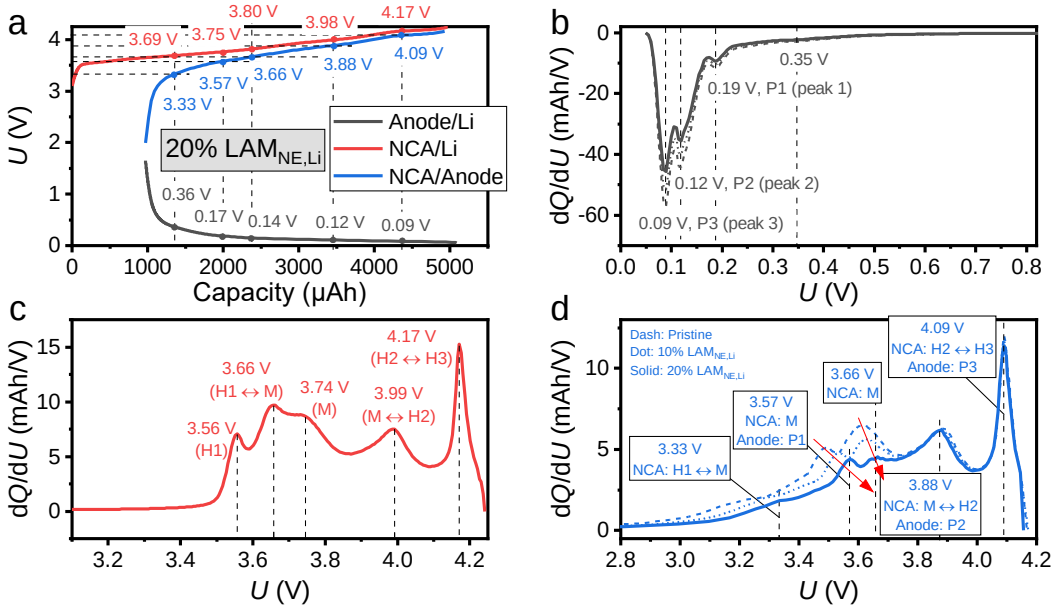
Supplementary Figure 11. d*Q*/d*U* plot with 20% loss of active materials of de-lithiated positive electrode ($LAM_{PE,de}$). **a**, Capacity vs. equilibrium potential (*U*) curves of anode/Li (grey), NCA/Li (red), and NCA/anode (blue). **b-d**, The d*Q*/d*U* plots of anode/Li (grey, **b**), NCA/Li (red, **c**), and NCA/anode (blue, **d**). The blue dots in **a** correspond to the peak positions in **d**, which help to identify the redox pairs in anode (**b**) and NCA (**c**) corresponding to the full cell (**d**). The dash, dot, and solid lines in **d** indicate the pristine, 10% $LAM_{PE,de}$, and 20% $LAM_{PE,de}$, respectively.



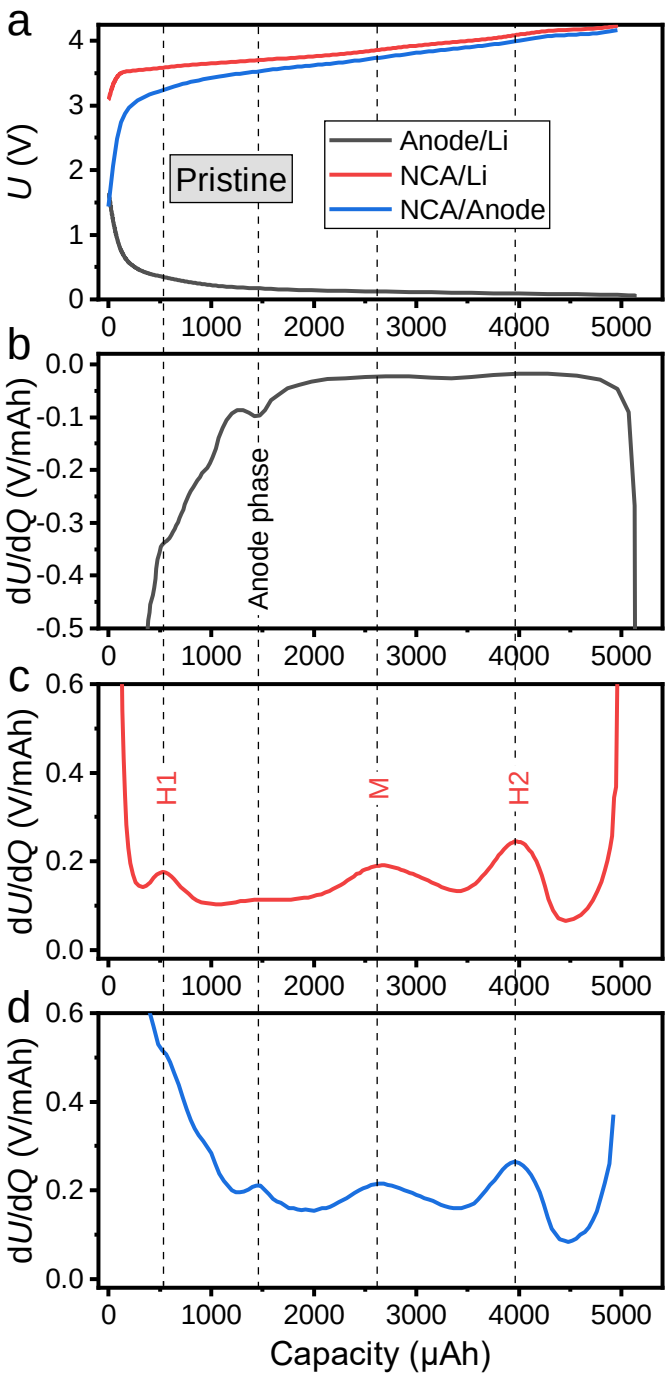
Supplementary Figure 12. d*Q*/d*U* plot with 20% loss of active materials of de-lithiated negative electrode ($LAM_{NE,de}$). **a**, Capacity vs. equilibrium potential (*U*) curves of anode/Li (grey), NCA/Li (red), and NCA/anode (blue). **b-d**, The d*Q*/d*U* plots of anode/Li (grey, **b**), NCA/Li (red, **c**), and NCA/anode (blue, **d**). The blue dots in **a** correspond to the peak positions in **d**, which help to identify the redox pairs in anode (**b**) and NCA (**c**) corresponding to the full cell (**d**). The dash, dot, and solid lines in **d** indicate the pristine, 10% $LAM_{NE,de}$, and 20% $LAM_{NE,de}$, respectively.



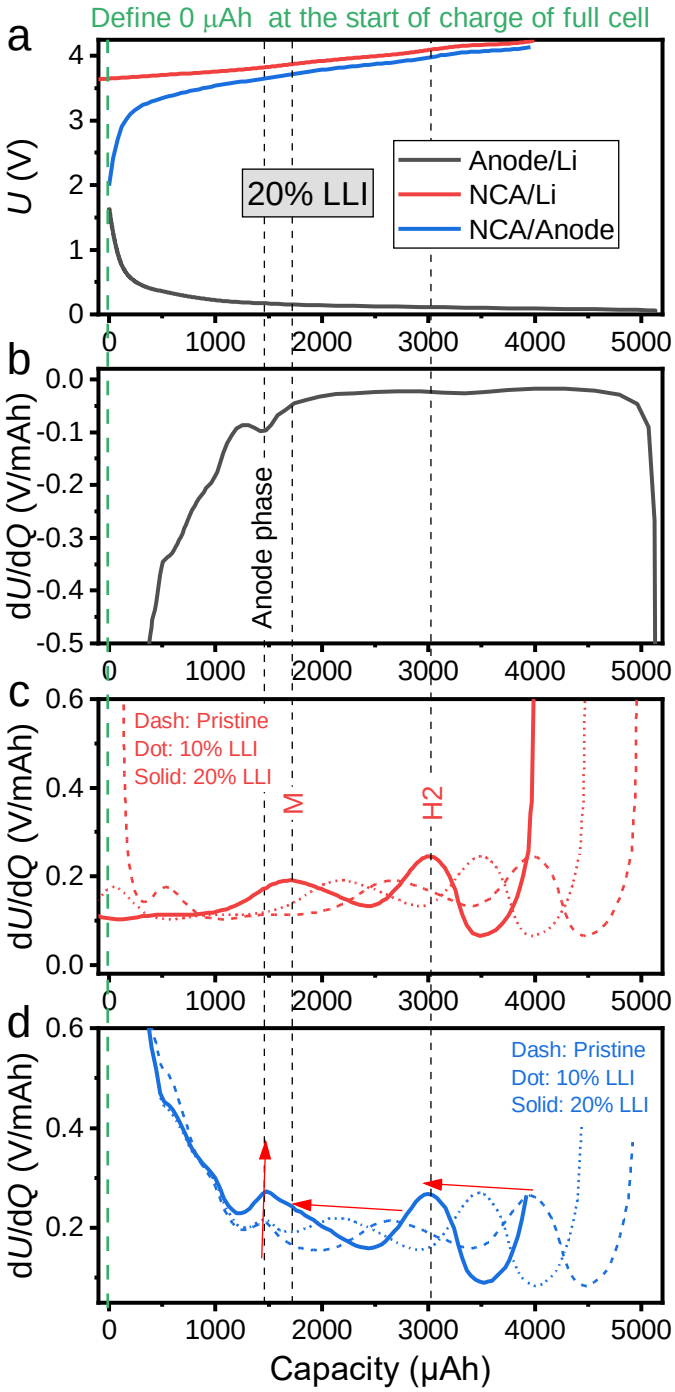
Supplementary Figure 13. dQ/dU plot with 20% loss of active materials of lithiated positive electrode (LAM$_{PE,Li}$). **a**, Capacity vs. equilibrium potential (U) curves of anode/Li (grey), NCA/Li (red), and NCA/anode (blue). **b-d**, The dQ/dU plots of anode/Li (grey, **b**), NCA/Li (red, **c**), and NCA/anode (blue, **d**). The blue dots in **a** correspond to the peak positions in **d**, which help to identify the redox pairs in anode (**b**) and NCA (**c**) corresponding to the full cell (**d**). The dash, dot, and solid lines in **d** indicate the pristine, 10% LAM$_{PE,Li}$, and 20% LAM$_{PE,Li}$, respectively.



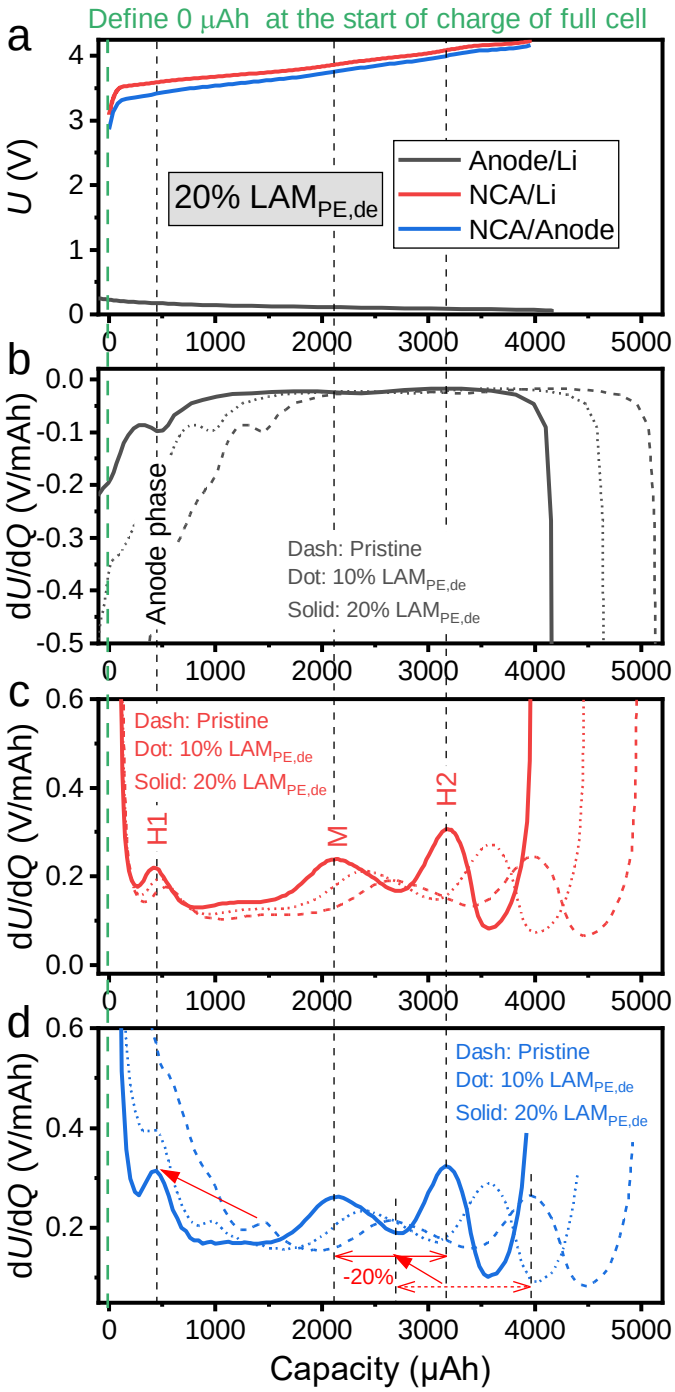
Supplementary Figure 14. dQ/dU plot with 20% loss of active materials of lithiated negative electrode (LAM$_{NE,Li}$). **a**, Capacity vs. equilibrium potential (U) curves of anode/Li (grey), NCA/Li (red), and NCA/anode (blue). **b-d**, The dQ/dU plots of anode/Li (grey, **b**), NCA/Li (red, **c**), and NCA/anode (blue, **d**). The blue dots in **a** correspond to the peak positions in **d**, which help to identify the redox pairs in anode (**b**) and NCA (**c**) corresponding to the full cell (**d**). The dash, dot, and solid lines in **d** indicate the pristine, 10% LAM$_{NE,Li}$, and 20% LAM$_{NE,Li}$, respectively.



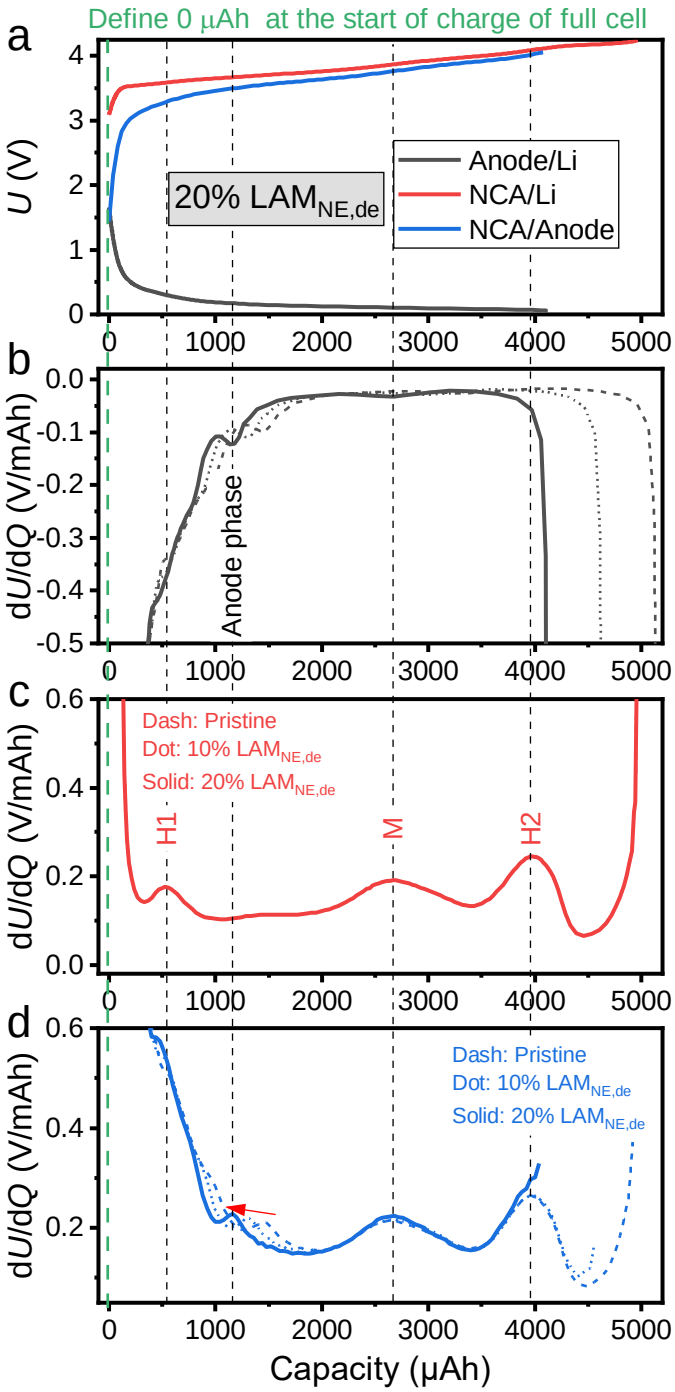
Supplementary Figure 15. d*U*/d*Q* plot at the pristine state. a, Equilibrium potential (*U*) vs. capacity curves of anode/Li (grey), NCA/Li (red), and NCA/anode (blue). **b-d**, The d*U*/d*Q* plots of anode/Li (grey, **b**), NCA/Li (red, **c**), and NCA/anode (blue, **d**). The vertical dash lines correspond to the peak positions in **d**, which help to identify the phases in anode (**b**) and NCA (**c**) corresponding to the full cell (**d**).



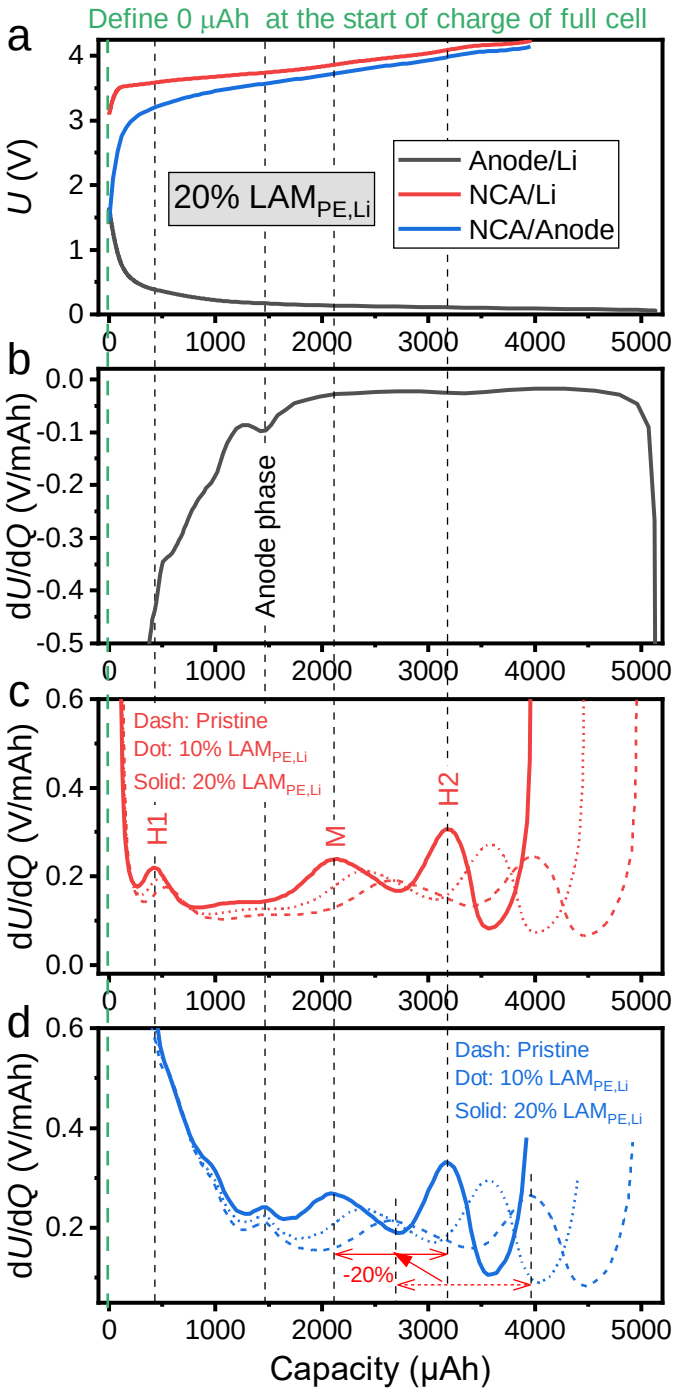
Supplementary Figure 16. d*U*/d*Q* plot with 20% loss of lithium inventory (LLI). a, Equilibrium potential (*U*) vs. capacity curves of anode/Li (grey), NCA/Li (red), and NCA/anode (blue). **b-d**, The d*U*/d*Q* plots of anode/Li (grey, **b**), NCA/Li (red, **c**), and NCA/anode (blue, **d**). The vertical dash lines correspond to the peak positions in **d**, which help to identify the phases in anode (**b**) and NCA (**c**) corresponding to the full cell (**d**). The dash, dot, and solid lines in **b-d** indicate the pristine, 10% LLI, and 20% LLI, respectively. The M peak of NCA is very close to the peak of anode phase, visually merging into one peak in full cell with 20% LLI.



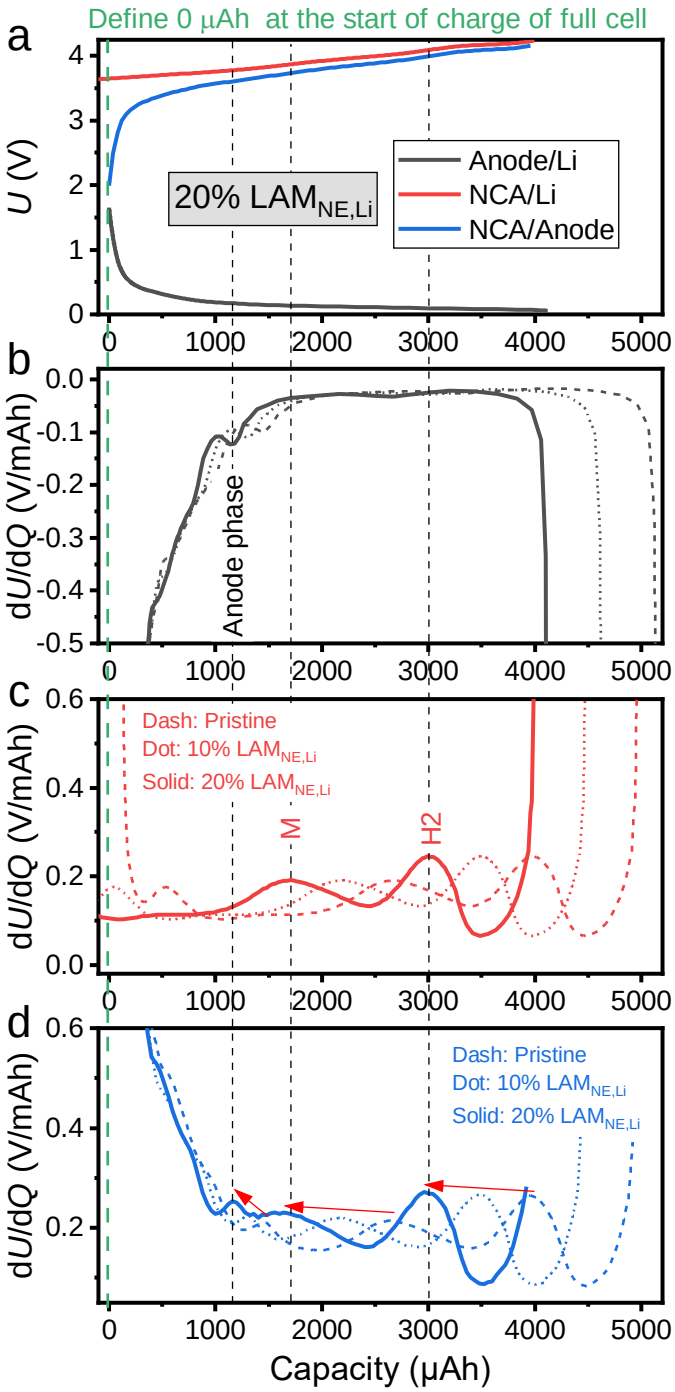
Supplementary Figure 17. d*U*/d*Q* plot with 20% loss of active materials of de-lithiated positive electrode ($LAM_{PE,de}$). a, Equilibrium potential (*U*) vs. capacity curves of anode/Li (grey), NCA/Li (red), and NCA/anode (blue). **b-d**, The d*U*/d*Q* plots of anode/Li (grey, **b**), NCA/Li (red, **c**), and NCA/anode (blue, **d**). The vertical dash lines correspond to the peak positions in **d**, which help to identify the phases in anode (**b**) and NCA (**c**) corresponding to the full cell (**d**). The dash, dot, and solid lines in **b-d** indicate the pristine, 10% $LAM_{PE,de}$, and 20% $LAM_{PE,de}$, respectively.



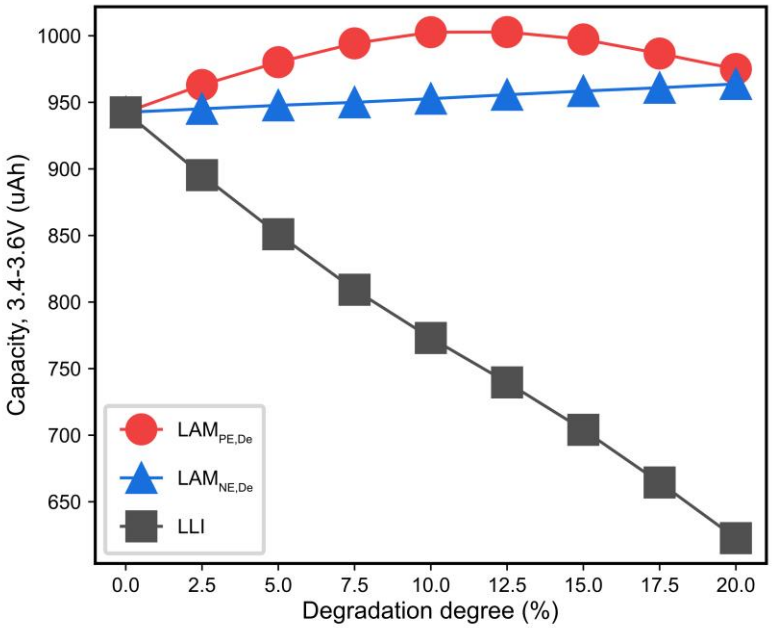
Supplementary Figure 18. d*U*/d*Q* plot with 20% loss of active materials of de-lithiated negative electrode ($LAM_{NE,de}$). **a**, Equilibrium potential (*U*) vs. capacity curves of anode/Li (grey), NCA/Li (red), and NCA/anode (blue). **b-d**, The d*U*/d*Q* plots of anode/Li (grey, **b**), NCA/Li (red, **c**), and NCA/anode (blue, **d**). The vertical dash lines correspond to the peak positions in **d**, which help to identify the phases in anode (**b**) and NCA (**c**) corresponding to the full cell (**d**). The dash, dot, and solid lines in **b-d** indicate the pristine, 10% $LAM_{NE,de}$, and 20% $LAM_{NE,de}$, respectively. Notably in **d**, the peak of anode phase left shifts without a considerable change in the amplitude. This phenomenon could serve as an important feature for the loss of $LAM_{NE,de}$.



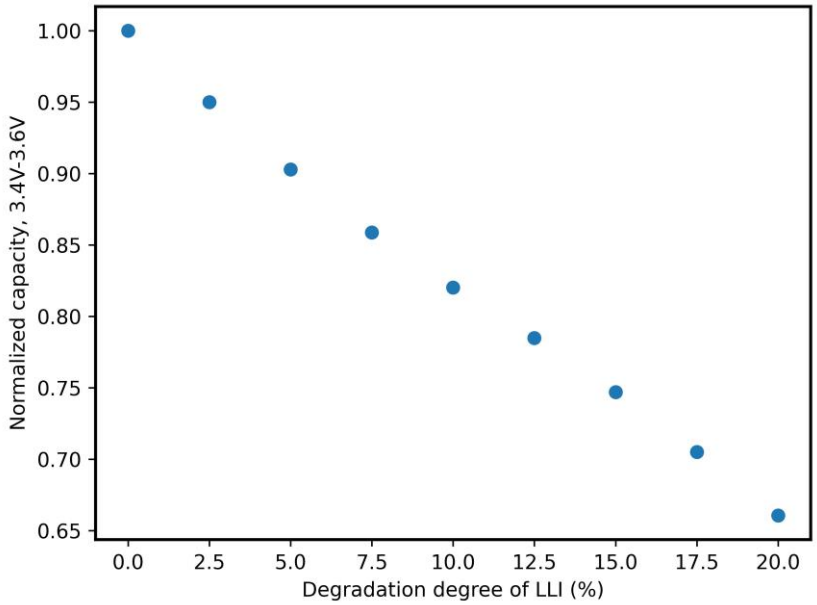
Supplementary Figure 19. d*U*/d*Q* plot with 20% loss of active materials of lithiated positive electrode ($LAM_{PE,Li}$). a, Equilibrium potential (*U*) vs. capacity curves of anode/Li (grey), NCA/Li (red), and NCA/anode (blue). **b-d**, The d*U*/d*Q* plots of anode/Li (grey, **b**), NCA/Li (red, **c**), and NCA/anode (blue, **d**). The vertical dash lines correspond to the peak positions in **d**, which help to identify the phases in anode (**b**) and NCA (**c**) corresponding to the full cell (**d**). The dash, dot, and solid lines in **b-d** indicate the pristine, 10% $LAM_{PE,Li}$, and 20% $LAM_{PE,Li}$, respectively.



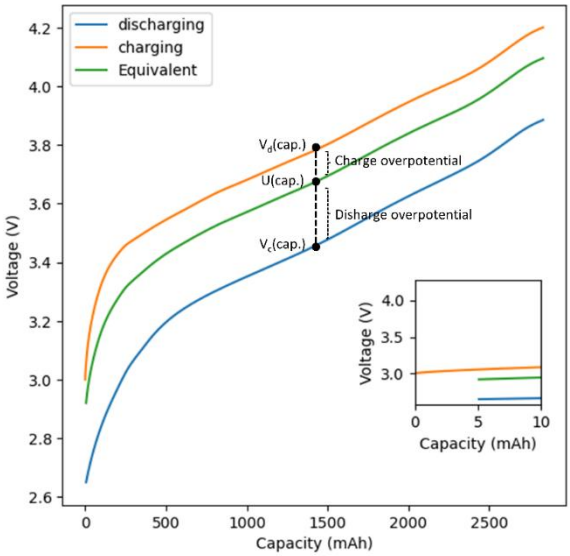
Supplementary Figure 20. d*U*/d*Q* plot with 20% loss of active materials of lithiated negative electrode (LAM$_{NE,Li}$). a, Equilibrium potential (*U*) vs. capacity curves of anode/Li (grey), NCA/Li (red), and NCA/anode (blue). **b-d**, The d*U*/d*Q* plots of anode/Li (grey, **b**), NCA/Li (red, **c**), and NCA/anode (blue, **d**). The vertical dash lines correspond to the peak positions in **d**, which help to identify the phases in anode (**b**) and NCA (**c**) corresponding to the full cell (**d**). The dash, dot, and solid lines in **b-d** indicate the pristine, 10% LAM$_{NE,Li}$, and 20% LAM$_{NE,Li}$, respectively.



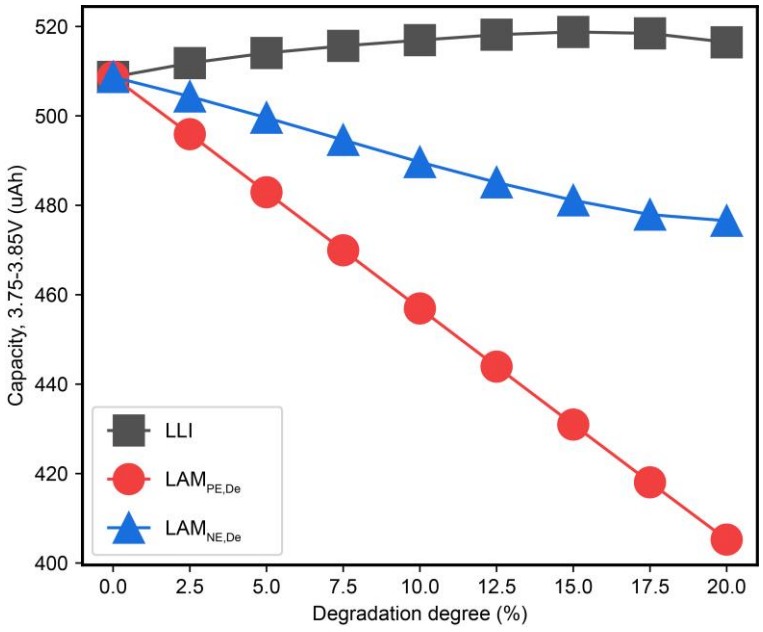
Supplementary Figure 21. Response of capacity between 3.4 and 3.6 V to different degradation modes. The capacity within this range is mainly influenced by the LLI, while $LAM_{PE,de}$ and $LAM_{NE,de}$ only minorly affect this capacity. Consequently, we can use this feature as a key descriptor for LLI.



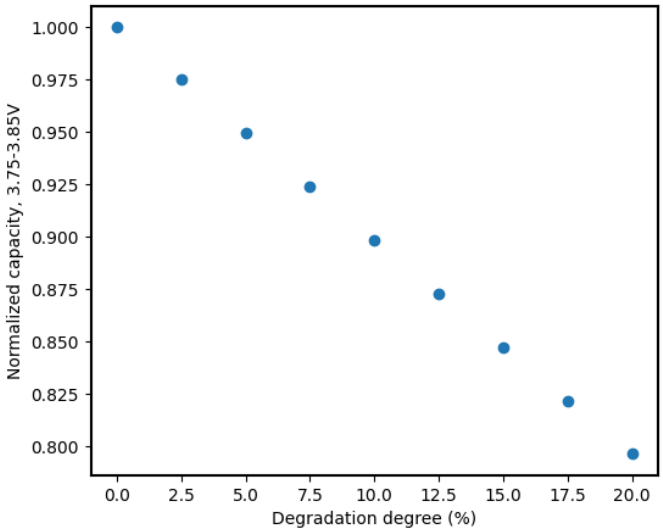
Supplementary Figure 22. The normalized capacity between 3.4V and 3.6V versus the degradation degree of LLI. The normalization is done by dividing each value by the capacity between 3.4V and 3.6V at the pristine state.



Supplementary Figure 23. Schematic diagram of calculating the equilibrium potential from the experimental data of voltage at relatively large currents. Due to the coulombic loss, the discharge capacity is slightly less than the charge capacity.



Supplementary Figure 24. Response of capacity between 3.75 and 3.85 V to different degradation modes. The capacity within this range is mainly influenced by the $LAM_{PE,De}$, while LLI and $LAM_{NE,de}$ only minorly affect this capacity. Consequently, we can use this feature as a key descriptor for $LAM_{PE,De}$.



Supplementary Figure 25. The normalized capacity between 3.75V and 3.85V versus the degradation degree of $LAM_{PE,De}$. The normalization is done by dividing each value by the capacity between 3.75V and 3.85V at the pristine state.

Supplementary References

1 Zhu, J. *et al.* Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation. *Nature Communications* 13, 2261 (2022). https://doi.org/10.1038/s41467-022-29837-w

2 Severson, K. A. *et al.* Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy* 4, 383-391 (2019). https://doi.org/10.1038/s41560-019-0356-8

3 Zhou, H. *et al.* Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. In *35th AAAI Conference on Artificial Intelligence, AAAI 2021, February 2, 2021 - February 9, 2021.* 11106-11115 (Association for the Advancement of Artificial Intelligence, 2021).

4 Li, W. *et al.* One-shot battery degradation trajectory prediction with deep learning. *Journal of Power Sources* 506 (2021). https://doi.org/10.1016/j.jpowsour.2021.230024

5 Loshchilov, I. & Hutter, F. Decoupled weight decay regularization. In *7th International Conference on Learning Representations, ICLR 2019, May 6, 2019 - May 9, 2019.* (ICLR, 2019).

6 Roman, D., Saxena, S., Robu, V., Pecht, M. & Flynn, D. Machine learning pipeline for battery state-of-health estimation. *Nature Machine Intelligence* 3, 447-456 (2021). https://doi.org/10.1038/s42256-021-00312-3

7 Saxena, A. *et al.* Metrics for evaluating performance of prognostic techniques. In *2008 International Conference on Prognostics and Health Management.* 1-17 (IEEE, 2008).

8 Birkl, C. R., Roberts, M. R., McTurk, E., Bruce, P. G. & Howey, D. A. Degradation diagnostics for lithium ion cells. *Journal of Power Sources* 341, 373-386 (2017). https://doi.org/10.1016/j.jpowsour.2016.12.011