nature

https://doi.org/10.1038/s41586-020-1994-5

Supplementary information

Closed-loop optimization of fast-charging protocols for batteries with machine learning

In the format provided by the authors and unedited

Peter M. Attia, Aditya Grover, Norman Jin, Kristen A. Severson, Todor M. Markov, Yang-Hung Liao, Michael H. Chen, Bryan Cheong, Nicholas Perkins, Zi Yang, Patrick K. Herring, Muratahan Aykol, Stephen J. Harris, Richard D. Braatz✉, Stefano Ermon✉ & William C. Chueh✉

Supplementary Information for

Closed-loop optimization of fast-charging protocols for batteries

Peter M. Attia[1†], Aditya Grover[2†], Norman Jin[1], Kristen A. Severson[3], Todor M. Markov[2], Yang-Hung Liao[1], Michael H. Chen[1], Bryan Cheong[1,2], Nicholas Perkins[1], Zi Yang[1], Patrick K. Herring[4], Muratahan Aykol[4], Stephen J. Harris[1,5], Richard D. Braatz[3*], Stefano Ermon[2*], William C. Chueh[1,6*]

† These authors contributed equally to this work

* Corresponding authors: wchueh@stanford.edu, ermon@cs.stanford.edu, braatz@mit.edu

1 Department of Materials Science and Engineering, Stanford University, Stanford, CA, USA

2 Department of Computer Science, Stanford University, Stanford, CA, USA

3 Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

4 Toyota Research Institute, Los Altos, CA, USA

5 Materials Science Division, Lawrence Berkeley National Lab, Berkeley, CA, USA

6 Applied Energy Division, SLAC National Accelerator Laboratory

Supplementary Tables

Supplementary Table 1. Features selected by the elastic net in the early outcome prediction model, along with the feature weights and feature scaling means and standard deviations. See Severson et al.[1] for additional details on feature definitions. Note that this model predicts the $\log_{10}$ of cycle life, not cycle life itself.

Feature	**Definition**	**Weight ($\times 10^{-3}$)**	**Scaling mean**	**Scaling standard deviation**
Initial discharge capacity	$Q_{\mathrm{discharge}}$, cycle 2	21.6	1.0736	0.0097
Max change in discharge capacity	$\max(Q_{\mathrm{discharge}}) - Q_{\mathrm{discharge}}$, cycle 2	−25.5	0.0054	0.0021
$\min(\Delta Q_{100\text{-}10}(V))$	$\log_{10}(\lvert\min(\Delta Q_{100\text{-}10}(V))\rvert)$	-21.3×10^{-3}	−1.3537	0.1830
$\mathrm{var}(\Delta Q_{100\text{-}10}(V))$	$\log_{10}(\lvert\mathrm{var}(\Delta Q_{100\text{-}10}(V))\rvert)$	−145	−3.6730	0.3758
$\mathrm{skew}(\Delta Q_{100\text{-}10}(V))$	$\log_{10}(\lvert\mathrm{skew}(\Delta Q_{100\text{-}10}(V))\rvert)$	12.9	−0.8553	0.4367

Supplementary Table 2. Comparison of initial discharge capacities (at cycle 2) of datasets used in Severson et al.[1] and the validation batch. Errors represent 95% confidence intervals. The packing date was 2015-09-26. A linear fit of mean initial discharge capacity vs time (assuming t=0 occurs at the pack date) reveals a slope of −1.2 mAh per 30 days (R^2=0.91).

Batch	**Date started**	**Mean initial discharge capacity (Ah)**	**Number of cells**
Train/primary test set 1	2017-05-12	1.078 ± 0.002	46
Train/primary test set 2	2017-06-30	1.071 ± 0.002	43
Secondary test set	2018-04-12	1.064 ± 0.002	46
oed_0 (this work)	2018-08-28	1.055 ± 0.004	46
oed_1 (this work)	2018-09-02	1.061 ± 0.002	46
oed_2 (this work)	2018-09-06	1.061 ± 0.002	48
oed_3 (this work)	2018-09-10	1.055 ± 0.003	48
Validation batch (this work)	2019-01-24	1.051 ± 0.003	45

Supplementary Table 3. Charging protocols and their associated CLO-estimated cycle life rankings, cycle lives, and one-sided standard deviations on cycle life at the end of batch 4. Each constant-current step (CC1, CC2, CC3, and CC4) comprises a 20% SOC window, such that CC1 ranges from 0 to 20% SOC, CC2 ranges from 20 to 40% SOC, etc. CC1, CC2, and CC3 are independent parameters, while CC4 is constrained by specifying that all protocols charge in the same total time (10 minutes) from 0% to 80% SOC.

Ranking	**CC1**	**CC2**	**CC3**	**CC4**	**CLO-estimated cycle life ($\mu_{4,i}$)**	**CLO cycle life std. deviation ($\sigma_{4,i}$)**
1	4.8	5.2	5.2	4.160	1185	78
2	5.2	5.2	4.8	4.160	1183	86
3	4.4	5.6	5.2	4.252	1174	76
4	4.8	5.2	4.8	4.457	1174	93
5	4	5.6	5.6	4.421	1164	93
6	4	5.6	5.2	4.707	1164	102
7	4.4	5.2	5.2	4.516	1162	94
8	4.4	5.6	5.6	4.017	1162	83
9	5.2	5.2	4.4	4.516	1160	116
10	5.2	5.2	5.2	3.900	1160	90
11	4.4	5.2	5.6	4.252	1159	88
12	4.8	5.2	5.6	3.935	1158	88
13	4.8	5.6	4.8	4.200	1158	76
14	4.8	5.6	5.2	3.935	1156	76
15	4.8	4.8	5.2	4.457	1149	106
16	4.4	5.6	4.8	4.563	1147	91
17	5.2	5.6	4.8	3.935	1144	84
18	4	6	5.2	4.457	1142	89
19	4.4	6	5.2	4.047	1140	82

20	5.2	4.8	4.8	4.457	1140	117
21	4	5.2	5.6	4.707	1137	117
22	5.2	4.8	5.2	4.160	1134	107
23	4.8	4.8	5.6	4.200	1133	104
24	4	6	5.6	4.200	1133	93
25	5.2	5.6	4.4	4.252	1130	93
26	4.8	5.6	5.6	3.733	1121	83
27	4.4	6	4.8	4.328	1120	83
28	4.4	6	5.6	3.834	1118	94
29	4.8	5.6	4.4	4.563	1117	98
30	4	6	4.8	4.800	1113	113
31	4.4	4.8	5.6	4.563	1110	116
32	5.2	5.2	5.6	3.702	1109	109
33	5.2	5.6	5.2	3.702	1107	86
34	5.6	5.2	4.4	4.252	1106	138
35	5.6	5.2	4.8	3.935	1106	127
36	4.8	6	4.8	4.000	1104	83
37	5.2	4.8	5.6	3.935	1103	131
38	3.6	6	5.6	4.755	1103	131
39	4.8	6	5.2	3.759	1097	84
40	5.2	5.6	4	4.707	1077	147
41	5.2	6	4.8	3.759	1071	93
42	5.6	4.8	4.8	4.200	1071	140
43	5.6	5.6	4.4	4.017	1070	123
44	5.6	4.8	4.4	4.563	1069	176
45	5.6	5.6	4.8	3.733	1068	120

46	5.6	5.2	4	4.707	1067	192
47	4.8	6	4.4	4.328	1067	94
48	5.6	5.2	5.2	3.702	1063	146
49	4.4	6	4.4	4.714	1062	114
50	5.2	6	4.4	4.047	1059	104
51	4.8	6	5.6	3.574	1058	97
52	5.6	4.8	5.2	3.935	1048	156
53	5.2	5.6	5.6	3.523	1047	100
54	4.8	4.4	5.6	4.563	1043	160
55	4.8	7	4.8	3.652	1039	114
56	5.2	7	4.8	3.450	1037	128
57	5.2	6	5.2	3.545	1033	93
58	4.8	7	5.2	3.450	1032	127
59	5.2	4.4	5.2	4.516	1032	150
60	5.6	5.6	4	4.421	1032	163
61	4.4	7	5.2	3.691	1022	143
62	5.2	7	4.4	3.691	1017	146
63	5.2	4.4	5.6	4.252	1017	163
64	5.6	5.6	5.2	3.523	1016	133
65	5.6	4.8	5.6	3.733	1016	190
66	5.2	6	4	4.457	1014	152
67	4.8	6	4	4.800	1013	158
68	5.6	5.2	5.6	3.523	1008	171
69	5.2	7	5.2	3.269	1008	148
70	5.6	6	4.8	3.574	1005	135
71	4.4	7	4.8	3.924	1005	133

72	4.8	7	4.4	3.924	1005	127
73	5.6	6	4.4	3.834	1003	147
74	5.6	7	4.8	3.294	998	191
75	4.4	7	5.6	3.513	996	160
76	6	5.2	4.4	4.047	994	182
77	6	4.8	4	4.800	989	253
78	6	4.8	4.4	4.328	987	208
79	5.6	4.4	4.8	4.563	987	151
80	4.8	7	5.6	3.294	987	162
81	4	7	5.2	4.030	986	179
82	6	5.2	4	4.457	985	213
83	5.6	7	4.4	3.513	982	188
84	6	5.2	4.8	3.759	978	188
85	5.2	6	5.6	3.381	978	116
86	4	7	5.6	3.818	977	188
87	5.6	4.4	5.2	4.252	975	150
88	5.2	7	4	4.030	971	190
89	6	4.8	4.8	4.000	971	189
90	5.6	6	4	4.200	967	180
91	4.8	7	4	4.308	965	168
92	5.6	7	5.2	3.129	965	197
93	5.6	4.4	5.6	4.017	962	189
94	4	7	4.8	4.308	961	166
95	5.6	6	5.2	3.381	961	140
96	4.4	7	4.4	4.239	959	133
97	3.6	7	5.2	4.537	955	236

98	3.6	7	5.6	4.271	955	242
99	6	5.6	4.4	3.834	954	164
100	5.6	5.6	5.6	3.360	954	155
101	6	4.4	4.4	4.714	953	215
102	5.2	7	5.6	3.129	953	183
103	6	5.6	4.8	3.574	947	180
104	6	4.8	5.2	3.759	944	210
105	6	7	4.8	3.170	941	245
106	6	5.6	4	4.200	940	186
107	6	4.4	4.8	4.328	939	165
108	6	5.2	5.2	3.545	938	212
109	5.6	6	3.6	4.755	935	240
110	5.2	7	3.6	4.537	933	245
111	7	4.8	4	4.308	933	213
112	7	4.4	4	4.690	932	222
113	5.6	7	4	3.818	932	195
114	6	5.6	3.6	4.755	929	244
115	6	4	4.8	4.800	927	169
116	4.4	7	4	4.690	926	170
117	6	4.4	5.2	4.047	926	180
118	6	7	4.4	3.372	926	225
119	7	3.6	5.6	4.271	925	271
120	6	4.8	5.6	3.574	922	242
121	6	7	5.2	3.017	918	244
122	4	7	4.4	4.690	918	169
123	6	4.4	5.6	3.834	918	227

124	5.2	4	5.6	4.707	918	206
125	7	3.6	5.2	4.537	916	234
126	6	4	5.2	4.457	916	163
127	5.6	7	5.6	3.000	916	209
128	6	6	4.8	3.429	914	199
129	6	3.6	5.6	4.755	912	247
130	7	4	5.6	3.818	911	232
131	7	4.4	4.4	4.239	911	194
132	7	4.8	4.4	3.924	909	201
133	6	6	4.4	3.667	909	194
134	5.6	4	5.2	4.707	908	160
135	6	5.6	5.2	3.381	908	205
136	6	4	5.6	4.200	907	211
137	5.6	6	5.6	3.231	906	168
138	7	4	5.2	4.030	905	177
139	7	5.2	3.6	4.537	904	252
140	6	5.2	5.6	3.381	901	235
141	5.6	4	5.6	4.421	899	187
142	7	4	4.4	4.690	898	192
143	7	4	4.8	4.308	898	167
144	7	4.4	4.8	3.924	893	172
145	7	5.2	4	4.030	892	220
146	6	7	5.6	2.897	888	244
147	5.6	7	3.6	4.271	888	226
148	6	6	5.2	3.250	887	213
149	6	6	4	4.000	887	209

150	7	4.4	5.2	3.691	881	158
151	6	7	4	3.652	880	194
152	6	6	3.6	4.500	879	250
153	7	4.4	5.6	3.513	878	199
154	7	4.8	4.8	3.652	876	183
155	7	5.2	4.4	3.691	874	205
156	6	5.6	5.6	3.231	869	227
157	8	3.6	5.6	3.969	862	259
158	6	6	5.6	3.111	855	237
159	8	4	4	4.800	855	222
160	8	4	5.6	3.574	854	264
161	8	4.4	5.6	3.306	851	262
162	7	5.6	3.6	4.271	849	233
163	8	4.8	5.6	3.111	848	250
164	8	5.2	5.6	2.963	846	265
165	6	7	3.6	4.065	845	203
166	7	4.8	5.2	3.450	845	156
167	7	5.2	4.8	3.450	844	191
168	8	3.6	5.2	4.197	837	226
169	8	5.6	5.6	2.847	837	294
170	8	4.4	4	4.328	836	218
171	8	3.6	4.8	4.500	833	242
172	7	5.6	4	3.818	831	214
173	7	4.8	5.6	3.294	831	174
174	8	4	4.4	4.328	829	229
175	8	5.2	5.2	3.089	828	218

176	8	4	5.2	3.759	826	232
177	8	4.8	5.2	3.250	826	203
178	8	4.4	5.2	3.463	825	228
179	8	4.8	3.6	4.500	822	251
180	7	5.6	4.4	3.513	822	207
181	8	5.6	5.2	2.963	821	266
182	8	4.4	4.4	3.940	818	212
183	8	4	4.8	4.000	817	235
184	7	7	3.6	3.706	817	287
185	8	4.4	4.8	3.667	814	221
186	7	6	3.6	4.065	813	217
187	8	4.8	4.8	3.429	813	186
188	8	5.2	4.8	3.250	811	189
189	7	5.6	4.8	3.294	810	215
190	7	5.2	5.2	3.269	809	181
191	7	7	4.4	3.122	809	237
192	7	7	4	3.360	808	270
193	8	4.8	4	4.000	805	203
194	8	4.8	4.4	3.667	805	185
195	8	5.6	4.8	3.111	802	238
196	8	6	5.6	2.754	797	303
197	7	7	4.8	2.947	796	195
198	7	5.2	5.6	3.129	794	191
199	7	6	4	3.652	791	205
200	7	5.6	5.2	3.129	790	233
201	7	6	5.6	2.897	789	284

202	7	6	4.8	3.170	789	235
203	7	6	4.4	3.372	788	208
204	8	5.2	4.4	3.463	785	175
205	7	5.6	5.6	3.000	785	248
206	8	6	5.2	2.862	784	290
207	7	6	5.2	3.017	784	266
208	8	6	4.8	3.000	775	273
209	7	7	5.2	2.814	774	176
210	7	7	5.6	2.710	773	199
211	8	5.2	3.6	4.197	767	213
212	8	5.6	4.4	3.306	766	209
213	8	5.2	4	3.759	762	173
214	8	6	4.4	3.181	745	249
215	8	7	3.6	3.476	733	202
216	8	5.6	4	3.574	728	178
217	8	5.6	3.6	3.969	725	193
218	8	6	3.6	3.789	712	228
219	8	6	4	3.429	710	224
220	8	7	4	3.170	709	176
221	8	7	4.4	2.957	706	173
222	8	7	4.8	2.800	684	172
223	8	7	5.6	2.585	650	200
224	8	7	5.2	2.680	648	174

Supplementary Table 4. Comparison of early outcome predictors. To contextualize the results, we included the fixed-mean baseline, which predicts a constant mean computed on the training set (held-out set is "primary test set" from Severson et al.[1]). See "Related work" section of Supplementary Discussion 2 for more details on the predictors proposed by Domhan et al.[2] and Klein et al.[3]

Method	**Training RMSE (cycles)**	**Held-out RMSE (cycles)**
Fixed-mean baseline	327	399
Domhan et al.[2]	681	756
Klein et al.[3]	110	190
"Discharge model" from Severson et al.[1] (similar to ML model used in this work)	76	86

Supplementary Table 5. Hyperparameter values used in the Bayesian optimization algorithm. The "Hyperparameter variable name in code" column refers to the variable name used in the BO code (closed_loop_oed.py). The hyperparameters are further detailed in the "Hyperparameter optimization" section of the Methods. The 2018-04-12 batch of Severson et al.[1] is used here because the cycling structure (i.e., rest steps, constant-voltage hold conditions, etc) is similar to that used in this work, although different charging protocols were tested.

Hyperparameter name	**Hyperparameter variable name in code**	**Value**	**Rationale**
Batch size	bsize	48	48-channel cycler
Budget	budget	4	~200 cells for optimization (4×48)
Standardization mean	standarization_mean	947	Mean cycle life of 2018-04-12 batch of Severson et al.[1]
Standardization standard deviation	standarization_std	164	Standard deviation of 2018-04-12 batch of Severson et al.[1]
Likelihood standard deviation	likelihood_std	164	Additive error (square root of the sum of the squares) of the sampling variation (126) and the prediction error (106) in 2018-04-12 batch of Severson et al.[1]
γ	gamma	1	Optimized via simulator
β_0	init_beta	5.0	Optimized via simulator
ε	epsilon	0.5	Optimized via simulator

Supplementary Table 6. Temperature dependence of SEI growth at C/10 as estimated from Figure 7 of Smith et al.[4], as used in the simulator. The data were manually extracted from the figure using WebPlotDigitizer. The slopes and intercepts were determined by linear fits of Q_{SEI} to $t^{0.5}$. The activation energy was determined by fitting the Arrhenius equation, using the slopes as approximations for *k*. The resulting activation energy is 0.122 eV.

Temperature (°C)	**Slope (mAh g^{-1} h$^{-0.5}$)**	**Intercept (mAh g^{-1})**
30	1.283	34.6
40	1.449	35.6
50	1.714	38.7

Supplementary Discussions

Supplementary Discussion 1. Alternate approaches for early outcome prediction and hyperparameter optimization

Early outcome predictor development and BO hyperparameter optimization are two critical components for closed-loop optimization. Both of these tasks require some initial data, which adds to the experimental requirements for the closed loop. Here we discuss our approach, as well as alternative approaches, to address these requirements.

For learning the early outcome predictor, we used a previously collected dataset[1] of 41 batteries cycled to failure as the training set. We note that this dataset contains reach information and can be used for early prediction of relevant outcomes of similar optimization tasks, such as different charging times and SOC windows; if multiple optimization objectives are to be considered for a new problem, we recommend collecting a dataset that spans the conditions of interest. For instance, all batteries in the training set were cycled with three-step or four-step charging protocols, as opposed to six-step. Additionally, while this training set was readily available, accurate predictive performance could likely be achieved with a smaller training set. Finally, previously collected datasets with different chemistries or cycling conditions could be leveraged for other battery optimization tasks using approaches such as transfer learning[5] and hybrid physics-data models.

For tuning hyperparameters of the BO component of our pipeline, we used both the training set for the early predictor and a simple physical simulator (see Methods). This simulator was originally designed for fundamental studies of heat transfer but was adapted to produce cycle life estimates as a function of charging protocol (see Physics-based simulator section of Methods). The only experimental input is the expected cycle life range. As such, the simulator does not accurately capture degradation during fast charging but provides reasonable estimates for hyperparameter optimization. Given that the simulator does not represent the true parameter space, the selected hyperparameter combination is likely suboptimal; however, we observe that the performance of the BO algorithm is relatively insensitive to the selected hyperparameters (Extended Data Figure 9), meaning that suboptimal hyperparameter combinations can still perform well.

Supplementary Discussion 2. Related work and possible modifications to closed-loop system

Related work review and comparison

The underlying optimal experimental design problem we seek to address is an instance of multi-fidelity black-box optimization. In black-box optimization, we only have sample access to an underlying function of interest. However, a single evaluation can itself be very expensive (in terms of time or resources), and hence low-fidelity signals that are noisy but easy-to-obtain are often used instead for efficient optimization.

The canonical use case considered in prior work is that of hyperparameter optimization for machine learning algorithms. Here, the goal is to find the optimal configuration of hyperparameters for learning a model on a target dataset. However, training a model with even a single configuration can be compute and time-intensive. To resolve this shortcoming, prominent examples of low-fidelity signals include (a) early prediction of learning curves (e.g., Domhan et al.[2]) and (b) testing a hyperparameter configuration on a subset of the data (e.g., Petrak et al.[6]). Multi-fidelity approaches (e.g., Van den Bosch et al.[7], Krueger et al.[8], Sabharwal et al.[9], and Sparks et al.[10]) use one or more of these low-fidelity signals for dynamically selecting which hyperparameter configuration to test next, e.g., *successive halving* routines[11] such as Hyperband[12], which periodically removes half of the worst performing configurations. We provide details on methods that predict learning curves, since that form of fidelity signal is similar to the one we consider. We further refer the reader to Hutter et al.[13] (Section 1.4) for an excellent and comprehensive survey on multi-fidelity black box optimization.

A learning curve tracks the performance of a machine learning algorithm (as measured via a suitable metric such as classification accuracy on a validation dataset) as a function of training time (as measured via number of iterations, gradient updates, etc.). These curves are typically increasing functions, and the performance metric saturates over time. Since observing the asymptote of a learning curve for any single hyperparameter configuration can be expensive, Domhan et al.[2] proposed to *predict* a learning curve after training the model only for a small number of iterations. The goal is to terminate the training procedure early if the asymptotic performance on a validation set is unlikely to beat the validation performance of the best known hyperparameter configuration so far. The predictor is parameterized using a mixture of functional forms (e.g., Weibull, vapor pressure, etc.) that have similar characteristics as learning curves (i.e., increasing and saturating with respect to the performance metric). Using standard Bayesian inference over the unknown parameters, the probability that the performance of the algorithm at a future iteration will exceed the best known performance so far can be estimated. Such estimates are then used within a downstream black-box optimization routine such as SMAC[14,15] or Tree Parzen Estimator[16] for the purpose of deciding whether to terminate training of a given hyperparameter configuration and utilize the remaining budget for testing other candidate hyperparameter configurations. Klein et al.[3] extend this approach via Bayesian neural networks to utilize additional data from past runs of related hyperparameter configurations. This permits learning more accurate predictors without having to observe a substantial portion of the learning curve.

To draw an analogy with the current work, we are interested in predicting the cycle life at which the capacity of a charging protocol drops to a predefined value (80% of initial capacity, or 0.88 Ah). In contrast to typical shapes of learning curves, (a) the capacity is decreasing as a function of cycle life and (b) the cycle life saturates with respect to the capacity (y-axis). That is, the capacity decreases slowly in the beginning until it hits a transition point. Beyond the transition point, the capacity quickly drops to 0.88 Ah. This transition point is unknown and occurs close to the final cycle life. As a result, predicting the final cycle life is challenging without having substantially cycled the battery to observe the transition point. To overcome this challenge, our early predictor[1] uses both information from prior capacity curves as well as additional features at every timestep (e.g., voltage and capacity). This method allows for highly accurate predictions only after the first 100 cycles and generalization to protocols outside our training dataset.

Empirically, we perform comparisons with the approaches of Domhan et al.[2] and Klein et al.[3] for early prediction of cycle lives based on capacity curves. Due to the aforementioned differences between the shapes of learning curves and capacity curves, naively plugging in capacity curve data within these approaches gives extremely poor estimates. Instead, we modified the problem by switching the x- and y-axes of a capacity curve: that is, the "rotated capacity curve" plots the cycle life on the y-axis against the capacity on the x-axis. Further, to match the monotonically increasing behavior of a learning curve, we consider an invertible transformation $x \rightarrow 1.1-x$ (since the nominal initial capacity is 1.1 Ah) for these curves. The empirical results post these transformations still lag our approach (Supplementary Table 4). We believe these results are largely because both Domhan et al.[2] and Klein et al.[3] do not use additional features (e.g., voltage) recorded *during* cycling of a battery, unlike our approach.

Further improvements to optimization approach

Our closed-loop system demonstrates both high efficiency and high accuracy on this parameter space given the specified budget. However, additional untested modifications could further improve the performance of the closed loop in some settings. We discuss some possible modifications to consider in future work here. Many of these are active areas of research within the broader research community working on black-box optimization.

Requiring fewer cycles for early outcome prediction. In this work, we use a single model that predicts the final cycle life after 100 cycles, similar to that developed by Severson et al.[1] Predictive models that achieve similar accuracy but require fewer cycles could further reduce the time per experiment.

Incorporating adaptive early outcome predictions (multi-fidelity optimization). In this work, all predictions are generated after a specified number of cycles has been generated (i.e., 100). An alternative approach, detailed in Grover et al.,[17] would dynamically incorporate predictions made at various cycle numbers. In general, the accuracy of early outcome predictions increases with additional cycling data. If a prediction at an early cycle number has low expected lifetime with sufficiently high confidence, the experiment could be terminated early, and a new experiment could be run in its place. In practice, however, "batching" experiments as we have done may have practical advantages. For example, because the early outcome predictions are quite sensitive to temperature, we sought to minimize the number of times that we opened the environmental testing chamber; given our single temperature chamber for all channels, continuous cell switching would lead to frequent temperature disruptions. Furthermore, the practical cell switching frequency is around once or twice per day; "continuous" cell switching would be enabled by robotic cell switching technologies.

An alternative multi-fidelity approach is multi-phase closed-loop optimization. For instance, we could consider two phases of closed-loop testing. The first would perform preliminary classification of charging protocols into low-cycle-life or high-cycle-life classes; Severson et al.[1] demonstrated high classification accuracy using only the first 5 cycles. After this initial rapid screening phase, the second phase would proceed on the high-cycle-life protocols using a regression model for cycle life prediction.

Incorporating more fine-grained uncertainties in early predictions. In our system, all predictions are treated as scalar values; prediction intervals are computed for each prediction but are only used to flag anomalous high-uncertainty predictions. However, we expect the magnitude of the prediction intervals to vary for different cells (due to differences in manufacturing) and as a function of charging protocol (since some protocols will induce more cycle life variation than others). In future work, we could update the distribution of CLO-estimated cycle lives using confidence intervals specific to an early outcome predictor as opposed to a uniform upper bound. This modification would account for the nonuniform confidence in each prediction.

Incorporating better batching strategies. We disallowed repeated runs of the same charging protocol within a given round for inducing more exploration. This does not exclude the possibility that the top-48 protocols in a batch are highly similar. In principle, we want to select protocols which score well on the explore-exploit frontier and are diverse to ensure more efficient coverage. Methods based on slice sampling (to filter out a good subset of protocols) followed by clustering techniques (to group similar protocols together) can be applied for a more diverse selection of charging protocols to test. On the other extreme, an even more advanced batching strategy for BO would also permit repeated testing of promising charging protocols even within the same round for protocols for faster reduction of uncertainty.

Studying a wider variety of acquisition functions. Our acquisition function to trade-off exploration-exploitation was based on the Upper Confidence Bound (UCB) criteria, which enjoys strong regret guarantees in many scenarios and is highly popular in the literature on multi-armed bandits. In future work, we would like to study the effect of other acquisition functions used in Bayesian optimization, such as expected improvement and Thompson sampling.

Specifying criteria for early termination of the closed loop. As illustrated in Figure 3 and Extended Data Figures 3a–c, CLO converges relatively quickly on the high-cycle life region of the parameter space. If the time and/or cells are at a premium, formal early-stopping criteria could be used to determine when acceptable cycle life estimates have been obtained in the fixed-budget setting, even before the specified budget is exhausted. In other words, the closed loop could be terminated early (i.e., before the specified budget is exhausted) if a performance objective was reached with sufficient confidence. For example, a threshold for the "change in cycle life" metric presented in Extended Data Figures 3a–c could be used, i.e., stop when the change in CLO-estimated cycle life with increasing rounds falls below some threshold. Confidence-based stopping criteria could also be used, i.e., stop when the bounds on the top protocol(s) fall below some threshold. Finally, the closed loop can be operated in the fixed-confidence setting (i.e., continue testing until confidence in the best protocol exceeds some threshold), as opposed to the fixed-budget setting used in this work (arguably more practical).

Continuous parameter space. In line with the prior algorithms that we extend in this work,[17,18] we optimized over a discretized parameter space of charging protocols. Using alternate algorithms based on continuous parameter spaces could lead to performance improvements, as the search space for continuous optimization approaches can subsume the protocols in the discretized space. However, for relatively smooth parameter spaces such as the one considered in this work, discrete optimization can perform well with sufficiently fine discretization.

Multi-objective closed-loop optimization. Fast-charging protocols need to balance two competing objectives: fast charging time and high cycle life. In this work, we fixed charging time and optimized for cycle life, since any choice of a multi-objective function would be arbitrary. However, instead of optimizing for cycle life with fixed charging time, we could optimize over a function that balanced both objectives. Ideally, multi-objective functions would be carefully developed with input from business units, consumers, and other stakeholders.

Supplementary Discussion 3. Degradation mechanism

Our results revealed that fast-charging protocols with nearly constant-current current profiles generally exhibited the highest lifetimes, while protocols with monotonically decreasing currents performed relatively poorly. The optimized charging protocols revealed in this work are unexpected since we expect lithium plating to be the dominant form of degradation during fast charging, and plating would be minimized by monotonically decreasing currents to avoid large negative overpotentials.[19–21] In general, this result highlights the need for data-driven approaches to this problem, as first-principles approaches fail to capture the dominant degradation modes. While work is ongoing to precisely determine the dominant degradation modes, we hypothesize that solid-electrolyte interphase (SEI) growth, exacerbated by high temperatures, is the dominant degradation mode during these extreme operating conditions:

- While lithium plating is generally expected at high charging rates, the high environmental temperature (30°C) is a temperature regime in which lithium plating is unfavorable.[22] The rate of heat generation during fast charging surpasses the convective cooling in the chamber, so the cell temperature increase during cycling is significant (Extended Data Figure 2). In general, side reaction rates increase exponentially with the inverse temperature from the Arrhenius relation.
- Resistive (Ohmic) heat generation is governed by I^2R (Çengel & Boles[23]). In Extended Data Figures 1e–f, we determined that the resistance is both Ohmic (linear overpotential vs. current relationship) and largely independent of SOC, particularly for minimally-cycled cells. For a constant R, I^2 is minimized via constant-current charging, which is consistent with our observation that the highest-lifetime protocols as identified by CLO are similar to

constant-current charging. This hypothesis is also loosely supported by Extended Data Figure 3e, in which a simple relationship between the sum of squared currents, sum(I^2) and the OED-estimated lifetimes yields a correlation coefficient of −0.84. We recognize, however, that temperature has a complex effect on battery degradation.

- Previous reports[24–29] have observed SEI growth leading to loss of active material via electrochemistry and microscopy on similar or identical LFP/graphite cylindrical cells. Loss of active graphite material is responsible for the rapid decrease in capacity observed in these cells during fast charging.[1,27] Thus, SEI growth is consistent with known degradation modes in these cells.

Supplementary Discussion 4. Bias in early outcome prediction

Figure 4c and Extended Data Figure 7 reveal a non-trivial bias in early outcome prediction for the validation protocols, in which nearly all final lifetimes fall well below their predictions. The mean predicted cycle lives exceed the mean final cycle lives by 145 cycles. If this bias is corrected for in the predictions, the root-mean-square error decreases from 173 cycles to 96 cycles, while the mean absolute percent error decreases from 20.1% to 9.46%, in line with the errors reported in Severson et al.[1] Importantly, the rankings of validation protocols are relatively unaffected (Extended Data Figure 7), implying a relatively consistent offset.

We attribute the source of this systematic bias to additional calendar aging, which lowers the initial discharge capacities and thus leads to larger predictions. A similar effect was presented in Supplementary Figures 6 and 7 and Supplementary Note 2 of Severson et al.[1] Calendar aging results in additional SEI growth, which consumes lithium inventory and reduces the capacity.[30,31] Supplementary Table 2 displays the mean initial discharge capacities of various datasets using the same cells. Although all cells have the same pack date (2015-09-26), the time between tests varies significantly (nearly 20 months between train/primary test set 1 and the validation batch), and thus so does the mean initial capacity. Given the feature weights and scalings in Supplementary Table 1, a change of 10 mAh in the initial discharge capacity leads to an additional 234 cycles predicted, which more than compensates for the bias observed here. Because the training set has uniform calendar life, this calendar life aging effect is not accounted for in the predictive model, and thus the cycle lives of the validation batch are overestimated.

Supplementary References

1. Severson, K. A. *et al.* Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy* **4**, 383–391 (2019).
2. Domhan, T., Springenberg, J. T. & Hutter, F. Speeding Up Automatic Hyperparameter Optimization of Deep Neural Networks by Extrapolation of Learning Curves. in *Proceedings of the 24th International Conference on Artificial Intelligence* 3460–3468 (AAAI Press, 2015).
3. Klein, A., Falkner, S., Springenberg, J. T. & Hutter, F. Learning curve prediction with Bayesian neural networks. in *Proceedings of the 2017 International Conference on Learning Representations* 1–16 (2017).
4. Smith, A. J., Burns, J. C., Zhao, X., Xiong, D. & Dahn, J. R. A High Precision Coulometry Study of the SEI Growth in Li/Graphite Cells. *J. Electrochem. Soc.* **158**, A447–A452 (2011).
5. Hutchinson, M. L. *et al.* Overcoming data scarcity with transfer learning. *arXiv:1711.05099 [cond-mat, stat]* (2017).
6. Petrak, J. *Fast Subsampling Performance Estimates for Classification Algorithm Selection*. 3–14 http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.28.3305&rep=rep1&type=pdf (2000).
7. Bosch, A. van den. Wrapped progressive sampling search for optimizing learning algorithm parameters. in *Proceedings of the Belgium-Netherlands Conference on Artificial Intelligence* 219–226 (2004).
8. Krueger, T., Panknin, D. & Braun, M. Fast Cross-validation via Sequential Testing. *J. Mach. Learn. Res.* **16**, 1103–1155 (2015).
9. Sabharwal, A., Samulowitz, H. & Tesauro, G. Selecting Near-Optimal Learners via Incremental Data Allocation. in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)* 2007–2015 (2016).
10. Sparks, E. R. *et al.* Automating Model Search for Large Scale Machine Learning. in *Proceedings of the Sixth ACM Symposium on Cloud Computing* 368–380 (ACM, 2015).

doi:10.1145/2806777.2806945.

11. Karnin, Z., Koren, T. & Somekh, O. Almost Optimal Exploration in Multi-armed Bandits. in *Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28* III–1238–III–1246 (JMLR.org, 2013).

12. Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A. & Talwalkar, A. Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization. *Journal of Machine Learning Research* **18**, 1–52 (2018).

13. *Automated Machine Learning: Methods, Systems, Challenges*. (Springer International Publishing, 2019).

14. Hutter, F., Hoos, H. H. & Leyton-Brown, K. Sequential Model-based Optimization for General Algorithm Configuration. in *Proceedings of the 5th International Conference on Learning and Intelligent Optimization* 507–523 (Springer-Verlag, 2011). doi:10.1007/978-3-642-25566-3_40.

15. Hutter, F., Xu, L., Hoos, H. H. & Leyton-Brown, K. Algorithm Runtime Prediction: Methods & Evaluation. in *Proceedings of the 24th International Conference on Artificial Intelligence* 4197–4201 (AAAI Press, 2015).

16. Bergstra, J. S., Bardenet, R., Bengio, Y. & Kégl, B. Algorithms for Hyper-Parameter Optimization. in *Advances in Neural Information Processing Systems 24* (eds. Shawe-Taylor, J., Zemel, R. S., Bartlett, P. L., Pereira, F. & Weinberger, K. Q.) 2546–2554 (Curran Associates, Inc., 2011).

17. Grover, A. *et al.* Best arm identification in multi-armed bandits with delayed feedback. in *Proceedings of the 21st International Conference on Artificial Intelligence and Statistics (AISTATS)* vol. 84 833–842 (2018).

18. Hoffman, M. W., Shahriari, B. & de Freitas, N. On correlation and budget constraints in model-based bandit optimization with application to automatic machine learning. in *Proceedings of the 17th International Conference on Artificial Intelligence and Statistics (AISTATS)* vol. 33 365–374 (2014).

19. Keil, P. & Jossen, A. Charging protocols for lithium-ion batteries and their impact on cycle life—

An experimental study with different 18650 high-power cells. *Journal of Energy Storage* **6**, 125–141 (2016).

20. Ahmed, S. *et al.* Enabling fast charging – A battery technology gap assessment. *Journal of Power Sources* **367**, 250–262 (2017).

21. Schindler, S., Bauer, M., Cheetamun, H. & Danzer, M. A. Fast charging of lithium-ion cells: Identification of aging-minimal current profiles using a design of experiment approach and a mechanistic degradation analysis. *Journal of Energy Storage* **19**, 364–378 (2018).

22. Waldmann, T., Wilka, M., Kasper, M., Fleischhammer, M. & Wohlfahrt-Mehrens, M. Temperature dependent ageing mechanisms in Lithium-ion batteries – A Post-Mortem study. *Journal of Power Sources* **262**, 129–135 (2014).

23. Çengel, Y. A. & Boles, M. A. *Thermodynamics: An Engineering Approach*. (McGraw-Hill Education, 2015).

24. Liu, P. *et al.* Aging Mechanisms of $LiFePO_4$ Batteries Deduced by Electrochemical and Structural Analyses. *J. Electrochem. Soc.* **157**, A499–A507 (2010).

25. Safari, M. & Delacourt, C. Aging of a Commercial Graphite/$LiFePO_4$ Cell. *J. Electrochem. Soc.* **158**, A1123–A1135 (2011).

26. Sarasketa-Zabala, E. *et al.* Understanding Lithium Inventory Loss and Sudden Performance Fade in Cylindrical Cells during Cycling with Deep-Discharge Steps. *J. Phys. Chem. C* **119**, 896–906 (2015).

27. Anseán, D. *et al.* Fast charging technique for high power $LiFePO_4$ batteries: A mechanistic analysis of aging. *Journal of Power Sources* **321**, 201–209 (2016).

28. Lewerenz, M., Marongiu, A., Warnecke, A. & Sauer, D. U. Differential voltage analysis as a tool for analyzing inhomogeneous aging: A case study for $LiFePO_4$|Graphite cylindrical cells. *Journal of Power Sources* **368**, 57–67 (2017).

29. Lewerenz, M., Warnecke, A. & Sauer, D. U. Post-mortem analysis on $LiFePO_4$|Graphite cells describing the evolution & composition of covering layer on anode and their impact on cell

performance. *Journal of Power Sources* **369**, 122–132 (2017).

30. Broussely, M. *et al.* Aging mechanism in Li ion cells and calendar life predictions. *Journal of Power Sources* **97–98**, 13–21 (2001).

31. Keil, P. *et al.* Calendar Aging of Lithium-Ion Batteries I. Impact of the Graphite Anode on Capacity Fade. *J. Electrochem. Soc.* **163**, A1872–A1880 (2016).