Sequential Bayesian Optimization for Accelerating the Design of Sodium Metal Battery Nucleation Layers

Adam Thelen[a,†], Murtaza Zohair[a,†], Jayanth Ramamurthy[a], Andrew Harkaway[a], Weimin Jiao[a], Mihir Ojha[a], Mahdi Ul Ishtiaque[a], Todd A. Kingston[a], Cary L. Pint[a,*], Chao Hu[a,b,*]

[a] Department of Mechanical Engineering, Iowa State University, Ames, IA, 50011
[b] Department of Mechanical Engineering, University of Connecticut, Storrs, CT, 06269
[†] Indicates authors contributed equally.
* Indicates corresponding authors. Emails: carypint@iastate.edu; chao.hu@uconn.edu

Abstract

Complex interactions between battery materials make design optimization difficult. In this work, we quickly and efficiently optimize three interacting design variables of a carbon nucleation layer used in sodium metal batteries. Unlike existing materials optimization approaches focusing on maximizing favorable material properties, we execute a multi-parameter materials optimization scheme that uses cell-level aging data as the target objective. We employ a Bayesian optimization algorithm that intelligently selects the nucleation layer designs to test in sequence, quickly finding the design that yields the highest lifetime. Results from simulation studies conducted after the fact indicate that a well-tuned Bayesian optimization algorithm can optimize the nucleation layer properties roughly five times faster than a random sampling approach and roughly two times faster than a poorly-tuned Bayesian algorithm. We then examine correlations on all 177 cells with different nucleation layers and use electrochemical impedance spectroscopy and imaging to propose the primary mechanism for nucleation layer performance, considering dead sodium as the dominant mode of capacity loss. The algorithms, hyperparameter tuning strategies, and post-optimization mechanism analysis used in this work broadly apply to other battery chemistries and electrode designs and are essential for quickly bringing metal battery performance on par with existing battery technologies.

Keywords: batteries; materials optimization; sodium metal; machine learning; Bayesian optimization

1. Introduction

The global push to electrification will require significantly more energy storage from batteries than is presently available. In order to meet these demands, new batteries must be developed and manufactured using more easily sourced materials [1]. However, researching and developing new battery materials is a costly and time-consuming process. Battery materials are a complex area of research where the materials often interact with one another making it difficult to explore the material design space in entirety [2], [3]. Most existing approaches to battery materials

development identify a desirable cell-level trait (e.g., high power capability), establish a link between the desired trait and the mechanisms that affect it (e.g., electrode thickness, charge transfer resistance), and then develop new materials focused on maximizing the material properties that affect the mechanism (e.g., electrode mass loading, particle size) [4], [5]. This approach is sometimes referred to as mechanistic modeling and models favorable material properties as a function of structure and composition using small, locally optimized models built using physical knowledge of the system [6]. The complex interactions between battery components often produce trade-offs where optimal performance may be achieved at a more median material property value, not at maximums or minimums [7]. However, studying a single mechanism or design variable in isolation does not guarantee good results when implemented into a battery cell with multiple other interacting components. Conversely, building a mechanistic model of a complex battery system with nonlinear interacting components is extremely time- and cost-intensive. In turn, if the relationship defining the favorable material property is not yet understood, accurate modeling for material design is entirely infeasible [8], [9].

To this end, machine learning-assisted experimental searching is being studied as a new method for rapidly developing battery materials and understanding the design space. Machine learning approaches largely ignore the underlying mechanisms and physics of material interactions and instead make inferences from data. In this approach, a machine learning surrogate model is trained using a small dataset of initial experiments to learn the relationship between the design variables (inputs) and the performance metric of interest (output) [10], [11]. Then, the fitted experiment surrogate model is used to guide future experiments in sequence, reaching an optimal design in as few iterations as possible. Machine learning assisted experimental searching has been used to discover and optimize various battery materials, manufacturing processes, and control protocols [12]–[18]. In [12], researchers developed a battery fast charging protocol to have minimal impact on cell lifetime by sequentially testing charging current profiles suggested by a Bayesian optimization algorithm. Similarly, researchers in [19] demonstrated a machine learning-guided autonomous robotic system used to search for non-aqueous lithium-ion battery electrolytes with high electrochemical stability. While the studies mentioned above are unique and promising, there is little work published that tackles the challenge of optimizing battery materials using cell-level performance metrics. It is challenging to close the gap between material-level optimization and cell-level performance since materials optimized in isolation are not guaranteed to produce excellent performance when implemented in cells. There is a great need to investigate materials optimization schemes that use cell cycling and aging data to formulate a target objective based on cell-level performance instead of component-level material properties.

Among all battery systems, sodium batteries are one of the most promising batteries beyond lithium-ion [20]. Specifically, the anode-free sodium battery has been demonstrated to exhibit energy density competitive with lithium batteries and leverage the native electrochemical stability of sodium metal in standard electrolyte media to achieve Coulombic efficiencies near or exceeding 99.9% [21]. However, a bottleneck toward design in this system relates to the design characteristics of the negative electrode where sodium metal is electroplated. The reversibility of the sodium electroplating reaction is highly dependent on the negative electrode surface properties due to its role in the sodium nucleation and growth process. Unfavorable conditions lead to uncontrolled SEI formation and the accumulation of electrochemically inactive or "dead" sodium [4], [5], [22], [23]. A number of strategies have been adopted to promote reversible sodium deposition, such as increasing surface area to reduce the local current density, using sodium wettable surfaces, and

using 3D electrodes that accommodate the large sodium expansion during plating [24]–[30]. Recent studies have shown that a thin nucleation layer that guides smooth, layer-by-layer depositions of sodium is critical to achieving optimal battery performance metrics [31], [32]. However, the first-principle insight behind the design of nucleation layers for sodium metal remains challenging due to the complexity of understanding nucleation at a liquid-solid interface. Consequently, it is difficult to translate the general design rules identified in the literature to designing nucleation layers when considering trade-offs such as increased surface area leading to a decrease in surface area but poorer dispersion in the electrode slurry [31], [33]. In turn, the sodium metal system is an ideal platform for applying design optimization at the electrode level, even though no current studies have focused on this promising system.

This work demonstrates an experimental campaign to concurrently optimize three related design variables of a carbon nucleation layer in sodium metal batteries that maximize cell-level cycling performance. The three design variables, namely the concentration of carbon black, the concentration of carbon nanotube additives, and the time spent oxidizing in air, are highly related and have been shown to affect plating reversibility significantly [26], [34], [35]. Unlike existing materials optimization works focusing on maximizing favorable material properties, we directly optimize cell-level performance using cell lifetime as the target objective. This direct approach ensures all inactive components in the cell are accounted for in the optimization process, ensuring a more generalizable result. Additionally, rather than selecting design variables that involve time-consuming measurements of material properties, we optimize process variables that have a clear relationship with the structure of our electrode. Doing so ensures our approach is more adaptable to settings with equipment and time limitations, such as in practical battery manufacturing settings. The well-tuned Bayesian optimization algorithm employed here can find a nucleation layer design with the highest lifetime five times faster than a random search approach. Moreover, we find that a poorly tuned Bayesian optimization algorithm is two times faster than the random searching approach, suggesting that implementing the algorithm can still generate significant time savings for most battery materials design projects.

Overall, our work demonstrates the advantages and effectiveness of directly using cycling performance as the objective of a multi-parameter battery component optimization study. Because a non-physics-based model is employed, this approach is easily adaptable to other battery components and cell chemistries. Furthermore, we provide valuable insights and strategies for others who wish to efficiently select algorithm hyperparameters before initiating a battery materials optimization campaign.

2. Methods

2.1. Nucleation Layer Preparation

All nucleation layers began as a slurry of thermally oxidized LITX 66 carbon black from Cabot Corp., multi-wall carbon nanotubes from Cheap Tubes, and PVDF binder using N-Methyl-2-pyrrolidone (NMP) as a solvent. The carbon nanotubes were roughly 10-20 nm in size. The carbon black (CB) and carbon nanotubes (CNT) concentrations were randomly sampled following the predefined ranges of 20–90 and 0–50 wt%, respectively, with 1-wt% increments. The upper

boundary on concentration is selected based on the upper limit of active material content that can be deposited using the doctor blading method, and the CNT concentration range is based on literature for CNT additives in battery electrodes [36]–[38]. Similarly, the thermal oxidation time was determined by randomly sampling values in the range of 0-8 hours with 1-hour increments. Thermal oxidation has been shown to affect the microstructure of carbon significantly [39], [40]. The range of values along with the minimum increment of each design variable are presented in **Table 1**.

The CB powder was thermally oxidized in air at 450 °C . After oxidation, the slurries were mixed and cast on conductive carbon-coated aluminum foil (MTI Corp.) using a doctor blade set at 150 μm. The cast slurries were then dried in a vacuum oven for 12 hours before constructing cells. All nucleation layers have ~1 mg/cm^2 mass loading and 30-40 μm thickness. This ensures that the variations in the nucleation layer properties, such as porosity, surface area, and conductivity, are functions of the design variables.

Table 1: Nucleation layer design variable space with minimum, maximum, and increment values specified for each variable.

Design Variable or Processing Condition	Minimum Value	Maximum Value	Increment
CB Concentration (w/w)	0.20	0.90	0.01
CNT Concentration (w/w)	0.00	0.50	0.01
PVdF Concentration (w/w)	0.10	0.40	0.01
Oxidation Time (h)	0.00	8.00	1.00

2.2. Cell Assembly, Cycling, and Characterization

All cells were made using 1M $NaPF_6$ mixed in diglyme electrolyte. Prior to mixing the electrolyte, the $NaPF_6$ salt was dried at 100°C for more than 12 hours in an Ar glovebox. To dry the diglyme, we used a 4A molecular sieve for more than 48 hours. The $NaPF_6$ salt is Strem brand and 99%, while the diglyme was Sigma-Aldrich 99.5%. Since the quality of the electrolyte can significantly impact cell performance, the electrolyte was only used to make cells up to one day after preparation. All of the half-cells were assembled using CR2032 stainless steel cases from MTI Corp. inside of an Ar glovebox. The half-cells use a sodium metal working electrode that is pressed onto a stainless-steel disc. All electrodes were punched from aluminum foil using a 1 cm diameter punch. We use two Celgrad 2325 separators soaked in electrolyte to separate the sodium metal from the nucleation layer. Post assembly, all cells were allowed to rest for 12 hours before beginning cycling.

Cycling was conducted at room temperature on a Neware BTS4000 battery tester. First, the cells are galvanostatically discharged to 0 V at a rate of 0.5 mA/cm^2. This activates the half-cell system. Cycling starts with a discharge (plating) of 0.25 mAh/cm^2 at a rate of 0.5 mA/cm^2. The discharge step takes roughly 30 minutes. Next, the cell is charged (stripping) until it reaches the upper cutoff voltage of 0.1 V. All electrochemical impedance spectroscopy (EIS) measurements were done using aa multichannel Metrohm Autolab tester at room temperature. EIS cycling

follows the standard cycling protocol, that is, discharge 0.25 mAh/cm^2 capacity at a rate of 0.5 mA/cm^2. We recorded the EIS spectra at cycles 1, 3, and 10, using a perturbation of 5 mV between 1 Hz and 10^5 Hz. All SEM images were performed with an FEI Inspect F50 SEM. Last, we used Micromeritics Instruments Corporation ASAP 2020 to measure nitrogen absorption.

3. Sequential Bayesian Optimization for Materials Design

Sequential Bayesian optimization is a popular method used for optimizing the designs of complex systems [7], [17], [18]. While any Bayesian model can be used, Gaussian process (GP) regression is the most popular model. The GP regression model estimates the performance metric as a function of the tunable design variables. After fitting the GP regression model to an initial dataset, new designs are selected and tested. In each iteration, their test results are added to the dataset to update the GP regression model. Efficient optimization is realized by using a carefully chosen acquisition function to intelligently select the next design to test [17]. Acquisition functions have been widely studied in literature across many different domains, with a select few remaining more popular than others because of their applicability to a wide range of design problems [17], [18], [41]–[43]. Acquisition functions come in various functional forms, each designed to balance the trade-off between exploring new regions in the design space with little information (*exploration*) or exploiting existing regions where the performance metric is already found to take on a favorable value (*exploitation*). In what follows, we discuss the GP regression model used in the sequential optimization framework and introduce the handful of acquisition functions compared in this study.

3.1. Gaussian Process Regression

We use GP regression to build a probabilistic machine learning model that approximates the performance metric (cell lifetime) over the entire design space [44]. Specifically, our effort involves training a GP regression model to learn the mapping from a vector of D input variables $\mathbf{x}$ to their corresponding target y. Similar to other supervised machine learning techniques, GP model training requires a training dataset organized as a pre-collected set of N_{train} input-output pairs, $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N_{\text{train}}}$. We can then assemble an $N_{\text{train}} \times D$ design matrix $\mathbf{X}_{\text{t}}$ by aggregating all N_{train} input vectors (column vectors), i.e., $\mathbf{X}_{\text{t}} = \left[\mathbf{x}_1, \ldots, \mathbf{x}_{N_{\text{train}}}\right]^{\text{T}}$, and an N_{train}-element target vector $\mathbf{y}_{\text{t}}$, i.e., $\mathbf{y}_{\text{t}} = \left[y_1, \ldots, y_{N_{\text{train}}}\right]^{\text{T}}$. In our problem formulation, the GP input vector $\mathbf{x}_i$ consists of three nucleation layer design variables (CB conc., CNT conc., and thermal oxidation time in hours) that affect the physical properties of the nucleation layer. The GP output variable y_i is the measured lifetime (cycles) of a half-cell built under the setting of $\mathbf{x}_i$. We assume the observation (y) of the underlying function (f) is corrupted by i.i.d. zero-mean Gaussian noise, giving rise to the Gaussian observation model, $y = f(\mathbf{x}) + \varepsilon$, where $f(\mathbf{x})$ is the underlying function we wish to approximate and $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ with σ_ε^2 being the noise variance. We then model the latent function, $f(\mathbf{x})$, using a GP, denoted as:

$$f(\mathbf{x}) = \mathcal{GP}\big(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\big) \tag{1}$$

where the mean function $m(\mathbf{x})$ and covariance function (also known as a kernel) $k(\mathbf{x}, \mathbf{x}')$ can be expressed as the following

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})] \tag{2}$$

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}\left[\big(f(\mathbf{x}) - m(\mathbf{x})\big)\big(f(\mathbf{x}') - m(\mathbf{x}')\big)^{\mathrm{T}}\right] \tag{3}$$

Our GP regression formulation assumes a zero-mean prior, i.e., $m(\mathbf{x}) = 0$, as is traditionally done, and let the covariance function model $f(\mathbf{x})$. The kernel we choose to use is the popular squared exponential kernel, defined as

$$k_{\mathrm{SE}}(\mathbf{x}, \mathbf{x}') = \sigma_{\mathrm{f}}^2 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2l^2}\right) \tag{4}$$

where σ_{f}^2 is the signal variance of the input, and l is the length scale. Fitting a GP regression model entails optimizing the two kernel parameters, σ_{f}^2 and l, and noise variance, σ_{ε}^2, which together form a set of hyperparameters. The exact version of the squared exponential kernel used in this work has a length scale term uniquely defined for each input variable in $\mathbf{x}$. This is more commonly called automatic relevance determination, or ARD for short. The GP regression model is fit to the data $\mathcal{D}$ by iteratively tuning the values of the three hyperparameters to minimize the negative log-likelihood of the observed data $\mathcal{D}$. This process is carried out automatically using the GPy package in Python.

After training, the fitted GP regression model can be used to make predictions at new, unseen test points $\mathbf{X}_* = \{\mathbf{x}_i^*\}_{i=1}^{N_{\mathrm{test}}}$ (untested nucleation layer designs we wish to estimate the lifetime at) for unknown observations $\mathbf{y}_* = \{y_i^*\}_{i=1}^{N_{\mathrm{test}}}$, where N_{test} is the number of test sample points. Our design problem aims at making predictions at $\mathbf{X}_*$, which may span the entire range of the design space outlined in **Table 1**, by interpolating between the training points. The first step in predicting the function values $\mathbf{f}_*$ at the test points $\mathbf{X}_*$ is to assemble the multivariate Gaussian prior as

$$\begin{bmatrix}\mathbf{y}_{\mathrm{t}} \\ \mathbf{f}_*\end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix}\mathbf{K}_{\mathbf{X}_{\mathrm{t}},\mathbf{X}_{\mathrm{t}}} + \sigma_{\varepsilon}^2\mathbf{I} & \mathbf{K}_{\mathbf{X}_{\mathrm{t}},\mathbf{X}_*} \\ \mathbf{K}_{\mathbf{X}_*,\mathbf{X}_{\mathrm{t}}} & \mathbf{K}_{\mathbf{X}_*,\mathbf{X}_*}\end{bmatrix}\right) \tag{5}$$

where $\mathbf{K}_{\mathbf{X}_{\mathrm{t}},\mathbf{X}_{\mathrm{t}}}$ is the covariance matrix consisting of covariances between function values at the training points $\mathbf{X}_{\mathrm{t}}$ and is calculated using **Eqn. 4**., $\mathbf{I}$ is the identity matrix, similarly, the matrix $\mathbf{K}_{\mathbf{X}_*,\mathbf{X}_*}$ is the matrix of covariances between the test points $\mathbf{X}_*$, and the off-diagonal matrices, $\mathbf{K}_{\mathbf{X}_*,\mathbf{X}_{\mathrm{t}}}$ and $\mathbf{K}_{\mathbf{X}_{\mathrm{t}},\mathbf{X}_*}$, are called the cross-covariances. As described in the original Gaussian observation model, $y = f(\mathbf{x}) + \varepsilon$, we assume all observations include an additive, zero mean Gaussian noise whose variance is σ_{ε}^2. Under this assumption, we must include the addition of the noise variance σ_{ε}^2 in each diagonal element of the training covariance matrix $\mathbf{K}_{\mathbf{X}_{\mathrm{t}},\mathbf{X}_{\mathrm{t}}}$, i.e., assuming $\mathbf{y}_{\mathrm{t}} \sim \mathcal{N}\big(\mathbf{0}, \mathbf{K}_{\mathbf{X}_{\mathrm{t}},\mathbf{X}_{\mathrm{t}}} + \sigma_{\varepsilon}^2\mathbf{I}\big)$.

The GP regression model can be used to make predictions at new test points $\mathbf{X}_*$ by calculating the conditional distribution $p(\mathbf{y}_*|\mathbf{X}_*, \mathcal{D})$, calculated as follows

$$p(\mathbf{y}_*|\mathbf{X}_*, \mathcal{D}) = \mathcal{N}(\mathbf{y}_*|\mathbf{m}_*, \boldsymbol{\Sigma}_*) \tag{6}$$

where

$$\mathbf{m}_* = \mathbf{K}_{\mathbf{X}_t,\mathbf{X}_*}^{\mathrm{T}}\left(\mathbf{K}_{\mathbf{X}_t,\mathbf{X}_t} + \sigma_\varepsilon^2\mathbf{I}\right)^{-1}\mathbf{y}_t \tag{7}$$

$$\mathbf{\Sigma}_* = \mathbf{K}_{\mathbf{X}_*,\mathbf{X}_*} - \mathbf{K}_{\mathbf{X}_t,\mathbf{X}_*}^{\mathrm{T}}\left(\mathbf{K}_{\mathbf{X}_t,\mathbf{X}_t} + \sigma_\varepsilon^2\mathbf{I}\right)^{-1}\mathbf{K}_{\mathbf{X}_t,\mathbf{X}_*} \tag{8}$$

where $\mathbf{m}_*$ and $\mathbf{\Sigma}_*$ are the predicted mean and covariance at the new test points $\mathbf{X}_*$, respectively.

3.2. Acquisition Functions

In any Bayesian optimization framework, the acquisition function is responsible for selecting the next design points to evaluate. The better the chosen points, the quicker and more effective the optimization process is. First, the GP regression model estimates the objective function over the entire design space. Then, the acquisition function is evaluated at each design point, and designs are ranked based on their acquisition function values. Acquisition functions have been the focus of many studies and vary in objective and complexity based on discipline.

Here, we investigate three commonly used acquisition functions: probability of improvement (PI), expected improvement (EI), and upper confidence bound (UCB). We selected these acquisition functions because they have probabilistic formulations that work well with the GP regression model's probabilistic output. The GP regression model predicts a Gaussian lifetime distribution, parameterized by the predicted mean and variance calculated using **Eqns. 7 and 8**. The selected acquisition functions' ability to leverage the predicted mean and variance is convenient and desirable for sequential searching because the predicted uncertainty provides additional information about the model's fit of the experimental results across the design space. Acquisition functions that consider both a predicted mean and variance have a more dynamic exploration/exploitation trade-off that is effective at sequential searching.

Through simulation studies, we compare the speed at which each acquisition function achieves the optimum nucleation layer design [18], [41]. Additionally, we examine methods of determining suitable hyperparameter values and provide insights useful to others who wish to implement Bayesian optimization algorithms for materials discovery in practice. In what follows, we introduce each acquisition function and discuss their formulation and implementation.

3.2.1. Probability of improvement

PI is a classic acquisition function widely used across many engineering disciplines [45]. The basic idea underlying PI is that it is most desirable to find a new design point that yields the highest probability of improvement above a predefined threshold [46]. The PI acquisition function leverages the GP regression model's predicted mean and variance and takes the following form

$$PI(\mathbf{x}) = P\left(\hat{f}(\mathbf{x}) \geq \tau\right) = \Phi\left(\frac{\mu_{\hat{f}}(\mathbf{x}) - \tau}{\sigma_{\hat{f}}(\mathbf{x})}\right) \tag{9}$$

where $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the standard normal distribution, $\mu_{\hat{f}}$ and $\sigma_{\hat{f}}$ are the GP model-predicted mean and standard deviation of function f at design point $\mathbf{x}$, and τ is a user-defined improvement threshold, typically defined as $\tau = \mu^+ + \delta$, with μ^+ being the maximum predicted lifetime from the GP regression model given the observed cells and the

hyperparameter δ controlling the exploration-exploitation trade-off. Later in this work, we examine selecting a suitable value of δ and provide insight into how one might choose a suitable value for δ before conducting an optimization campaign.

3.2.2. Expected improvement

Another classical acquisition function is EI. The EI acquisition function extends the definition of the traditional PI acquisition function to include an additional term that considers the magnitude of improvement a new design point x might yield [47]. The EI acquisition function has the following form

$$EI(\mathbf{x}) = E\left[\left(\hat{f} - \tau\right)H\left(\hat{f} - \tau\right)\right] = \begin{cases} \sigma(\mathbf{x})\phi\left(\dfrac{\mu_{\hat{f}}(\mathbf{x}) - \tau}{\sigma_{\hat{f}}(\mathbf{x})}\right) + \left(\mu_{\hat{f}}(\mathbf{x}) - \tau\right)\Phi\left(\dfrac{\mu_{\hat{f}}(\mathbf{x}) - \tau}{\sigma_{\hat{f}}(\mathbf{x})}\right) & \sigma_{\hat{f}}(\mathbf{x}) > 0 \\ 0 & \sigma_{\hat{f}}(\mathbf{x}) = 0 \end{cases} \tag{10}$$

where $\phi(\cdot)$ denotes the probability density function (PDF) of the standard normal distribution, and $H(\cdot)$ is the standard Heaviside step function ($H\left(\hat{f} - \tau\right)$ equals one if $\hat{f} - \tau > 0$, and 0 otherwise). Similar to the PI acquisition function, the EI acquisition function analyzes the design space considering a minimum improvement threshold $\tau = \mu^+ + \delta$.

3.2.3. Upper confidence bound

The third acquisition function we analyze in this study is the upper confidence bound (UCB) function. The UCB acquisition function takes the following form

$$UCB(\mathbf{x}) = \mu_{\hat{f}}(\mathbf{x}) + \beta\sigma_{\hat{f}}(\mathbf{x}) \tag{11}$$

where β is a hyperparameter used to control the exploration/exploitation trade-off. At first, when few design points have been tested, it is generally more advantageous to explore the design space by selecting points in regions of highest uncertainty. Later, as the optimization continues, it is better to select design points near existing points that are observed to have high objective function values (cell designs with long lifetimes). The hyperparameter β takes the form $\beta_k = \beta_0\epsilon^k$, where β_0 and ϵ are parameters that determine the decay rate over each iteration k. Here, β_k decreases over each iteration to slowly transition from exploring the design space to exploiting existing designs.

4. Results and Discussion

4.1. Dataset Creation

To study the variables of conducting a materials optimization study in practice, we constructed a dataset in advance and examined it through optimization simulations. We randomly generated 32 unique nucleation layer designs to build the dataset by randomly sampling the design variable space outlined in **Table 1** using the Latin Hypercube sampling technique. We then constructed two sodium metal half-cells for each of the 32 nucleation layer designs, a total of 64 cells. Sodium metal was used as the counter/reference electrode in the half-cells along with 1M $NaPF_6$ diethylene glycol dimethyl ether (diglyme) electrolyte. Each of the 32 nucleation layer designs is visualized in **Fig. 1d,** where each point in the 3D space represents a unique set of design variable values and is colored with respect to the maximum lifetime of the pair of cells.

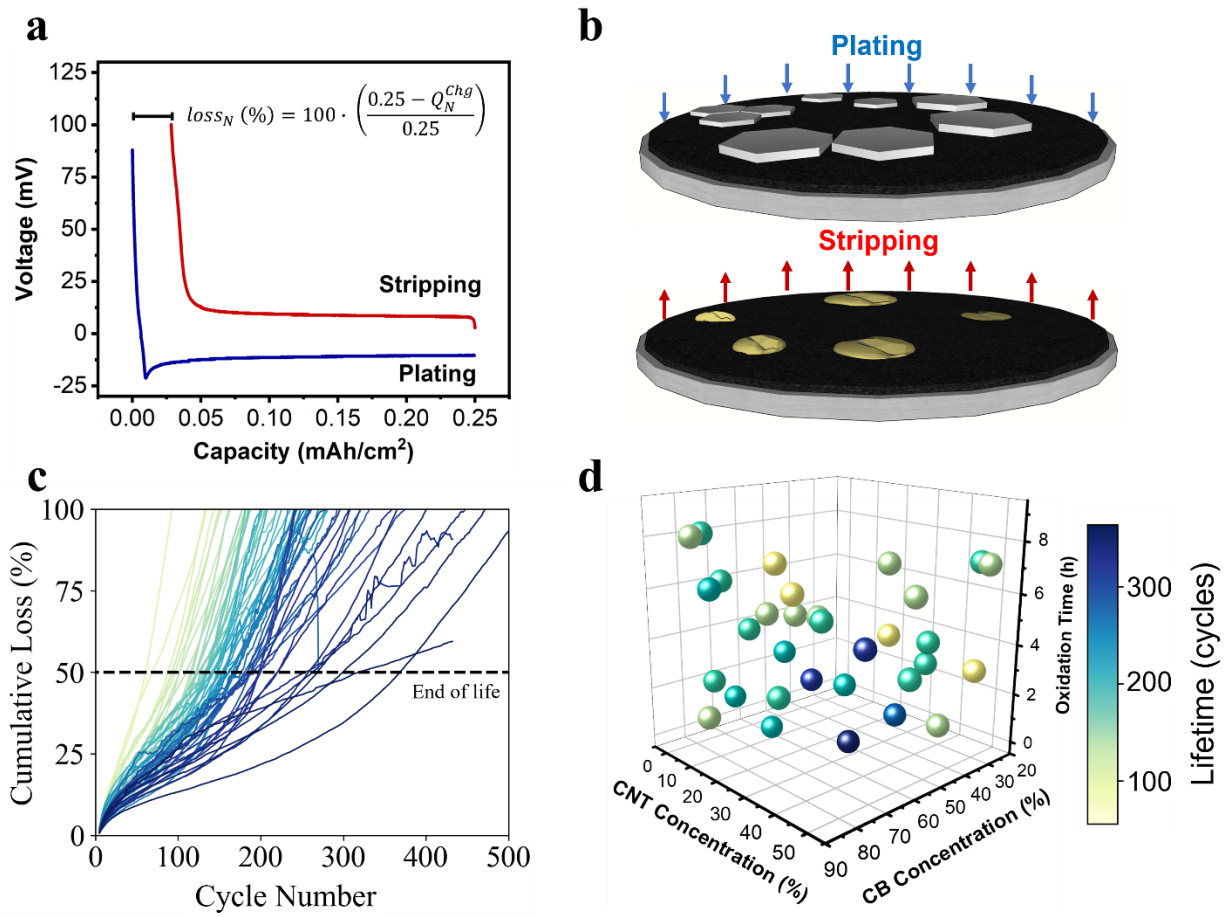

Figure 1: (a) The sodium metal plating and stripping process for a representative half-cell, visualized as the charge and discharge curves. The cycling loss is calculated as the difference between the plating and stripping capacities. (b) An illustration of the sodium metal plating and stripping process taking place of the carbon-based nucleation layer. (c) The cumulative loss of capacity as a function of cycle number for all cells. The dashed line at 50% represents the

threshold used to determine the end of life. (d) A 3D scatterplot of the 32 unique nucleation layer designs colored with respect to the maximum lifetime of the pair of cells.

We quantitatively assess the cell-level performance of the different nucleation layer designs by analyzing their cycling performance. The cells were galvanostatically cycled at 0.5 mA/cm^2 for roughly 500 cycles. Reporting CE is common method of evaluating cell performance, and in this work, we use CE to evaluate total charge losses. We first calculate the relative capacity loss at each cycle N using the following

$$loss_N\ (\%) = 100 \cdot \left(\frac{0.25 - Q_N^{\mathrm{Chg}}}{0.25} \right) \tag{12}$$

where Q_N^{Chg} is the charge capacity at the N-th cycle, and the constant 0.25 represents the fixed discharge capacity of 0.25 mAh/cm^2 that is constant for all cycles. Since discharge capacity (plating) is constant each cycle, all capacity losses arise from charging (stripping). We then calculate a cell's relative capacity-fade trajectory by cumulatively summing the loss at each cycle (see **Fig. 1c**). Cells are then compared in terms of their total lifetime, calculated as the number of cycles until their relative charge capacity loss exceeds the 50% threshold. We selected 50% for the threshold because it struck a balance between capturing long-term cycling stability and limiting the time spent cycling cells to no more than eight weeks.

In **Figure 1a**, we visualize the sodium metal plating and stripping process for a representative half-cell. Sodium metal is plated onto the nucleation layer surface during the discharge step until a capacity of 0.25 mAh/cm^2 has been reached. Then, the sodium metal is stripped off the nucleation layer until an upper voltage limit of 100 mV is reached. This ensures that he stripping reaction can occur, but that we avoid completely desodiating the carbon nucleation layer. We quantify the loss of sodium each cycle as the difference between the charge and discharge capacities, calculated as a percentage using **Eqn. 12.** Each cycle, internal side reactions between the metallic sodium and the other components in the cell lead to dead sodium formation, captured as the small difference in capacity, and best quantified through the cell's CE and lifetime. The imperfect plating and stripping process is illustrated in **Fig. 1b**, where dead sodium can be seen remaining on the nucleation layer during stripping.

A cell's lifetime is determined by its rate of dead sodium formation over many cycles. To determine a cell's lifetime, we cumulative sum the loss from each cycle and record the cell's lifetime as the cycle at which the loss exceeds 50%. This threshold roughly corresponds to a total anode-free full-cell capacity fade of 25%. The loss trajectories for all cells over the first 500 cycles are plotted in **Fig. 1c**. We observe that there is some cycle-to-cycle variability indicated by the bumpy trajectories, likely due to the instability of the plating and stripping process. We also observe that the initial capacity loss for most cells is steep, indicating increased dead sodium formation during early cycles, likely caused by SEI formation. After the early cycles, cycling appears to stabilize for most cells. Then, towards the end of life, many cells experience increasing loss of sodium each cycle until they reach the end-of-life threshold.

4.2. Sequential Bayesian Optimization Study

An overview of the sequential Bayesian optimization methodology is shown in **Fig. 2**. The sequential optimization process begins in **Fig. 2a**, where we identify three quantifiable design variables that can be manipulated to affect the material properties of the nucleation layers in the sodium metal half-cells. Next, an initial batch of samples must be tested to fit the GP regression model and approximate the objective function over the entire design space. A general guideline is to test just enough samples to cover the range of each design variable. We explore the impact of the initial number of samples on the algorithm's performance later in this section. After the initial samples are tested and the lifetimes determined (**Fig. 2b**), a GP regression model can be fit to the data to estimate cell lifetime as a function of its three nucleation layer design variables (**Fig. 2c**). Next, the GP regression model is used to estimate cell lifetime over the entire design space. All the possible nucleation layer designs are then ranked using one of the three acquisition functions outlined in **Sec. 3.2.** (**Fig. 2d**). The process starts again by building nucleation layers for new cells using the design variable values suggested by the acquisition function (**Fig. 2a**). The sequential process of building cells and evaluating their performance repeats until a suitable optimum is achieved.

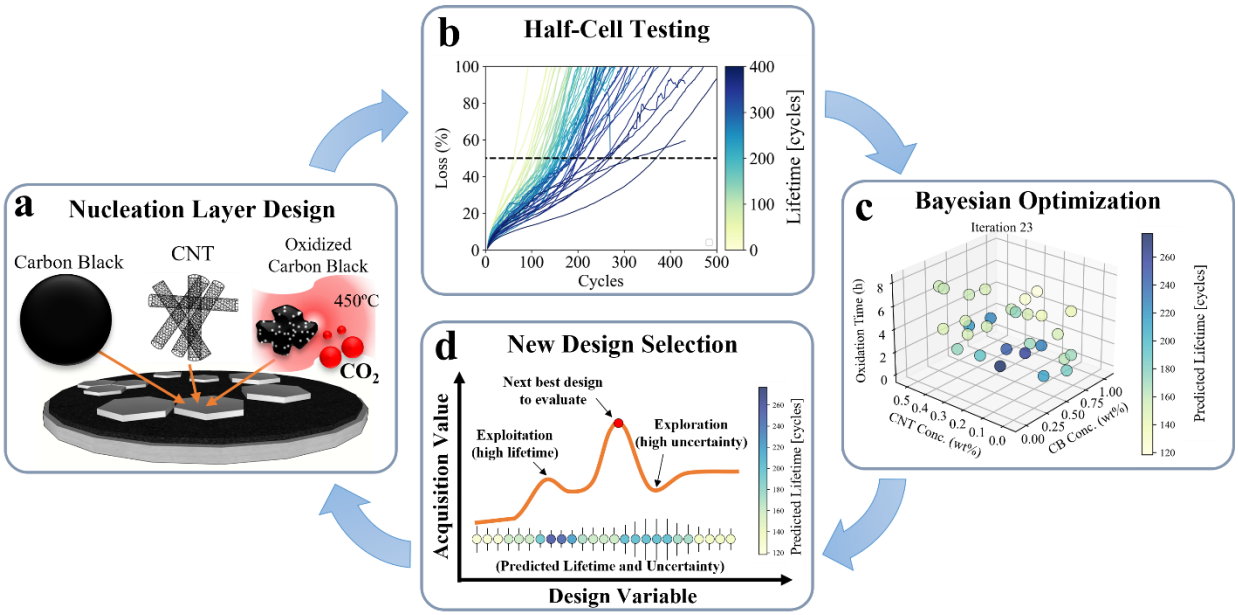

Figure 2: Overview of the closed-loop optimization framework used to optimize the design of the nucleation layer in a sodium metal half-cell. (a) Sodium half-cells with different nucleation layer designs are constructed by altering the three design variables: CB concentration, CNT concentration, and oxidation time. (b) half-cell total capacity loss as a function of the cycle number. (c) 3D design variable space where the GP regression model has estimated the lifetime of cells as a function of their design variable values. (d) visualization of the acquisition function used to select the next design points to test calculated over the entire design variable space.

4.2.1. Simulation Results

Sequential optimization simulations are carried out to analyze the impact of different factors on the speed and behavior of the optimization algorithm. We examine the sensitivity of the algorithm to the number of initial data points, the acquisition functions used, and the values of the acquisition function hyperparameters. Optimization simulations are run by first selecting a small number of initial samples to fit the GP regression model. Besides the study that investigates the initial number of starting datapoints, we randomly choose six samples as a starting dataset for all other tests. We then fit the GP regression model to the starting dataset and estimate the lifetime over the entire design variable range. Then, the acquisition function is used to select the next nucleation layer design to test. However, since we have evaluated the GP regression model over the entire range of design variable values, the acquisition function may select a design point at which we have not tested a cell. In this case, we calculate the Euclidean distance between the chosen design point and all remaining unobserved data points in the dataset. The cell closest to the chosen design point is selected and added to the observed dataset for the next iteration. While this approximation is not perfect, it is good enough to enable us to run more complex optimization studies without continuously testing cells. The process of fitting the GP regression model, evaluating the acquisition function, and adding a new point to the dataset is repeated until no data points remain. Since the initial batch of cells was randomly selected, all sequential optimization studies are repeated 100 times, and results are reported as averages of the multiple runs. Furthermore, each acquisition function uses an optimized set of hyperparameters.

Figures 3a, 3b, and **3c** summarize the simulation results characterizing the performance of each acquisition function as compared to a random sampling approach which randomly selects the next design point to test. The significance of using random sampling as a baseline optimization method is two-fold. First, random sampling is chosen because it is similar to the traditional design of experiments (DOE) optimization approach that aims to discover the global maximum by testing points throughout the entire design space. Both random sampling and DOE cannot adapt to the results of the experiments as they are performed. For DOE, the test points are predetermined prior to conducting any tests, and for random sampling, the next test points are chosen randomly each iteration to fill the design space. Second, random sampling is chosen as a baseline optimization approach because it is a viable method for solving high-dimensional optimization problems. While this work only considers three design variables, batteries are complex systems, and it is entirely possible that new battery materials optimization problems may be conducted with ten or more design variables, posing a greater challenge. Additionally, random sampling has been shown to be an effective method for high-dimensional optimization in other engineering disciplines [17], [18], [43].

In **Fig. 3a**, we plot the maximum achieved cell lifetime vs. the number of experimental iterations. This plot is used to assess how quickly each algorithm leads to a set of design variables exhibiting improved half-cell lifetime over the previous iteration. Each of the curves is strictly monotonically increasing because any cells tested that do not perform better than the current best cell are not of interest and thus reflected in the plot as no change in maximum lifetime over the previous iteration. From the results in **Fig. 3a,** we can see that all three tested acquisition functions reach the optimal nucleation layer design much quicker than the random sampling approach, roughly 10–15 iterations sooner on average. In particular, the PI and EI acquisition functions reach

the optimal design in approximately six to seven iterations, while the UCB acquisition function averages closer to 15 iterations.

This result suggests two key findings. First, the PI and EI acquisition functions performed similarly because they both considered the probability of improvement in their formulation. If the probability of improvement term were not dominant, we would expect the EI acquisition function to perform significantly differently; however, it did not. This result suggests that the excellent performance of both PI and EI is due to their formulation that considers the probability of improvement a new point may yield if added to the dataset. Second, the significant difference between the UCB acquisition function and the other two is also due to differences in formulation. The UCB acquisition function directly considers the magnitude of objective function improvement over a threshold, where the threshold is a scaled form of the predicted variance. If the predicted variance is large and uniform over the entire design space, the UCB acquisition function will struggle to pick good designs. This happens because the large variance will mask the smaller changes in the predicted lifetimes and alter the acquisition function's design ranking. In contrast, the probability of improvement term used in PI and EI normalizes the predicted mean by the predicted variance. This normalization helps to filter out noise in the experimental results.

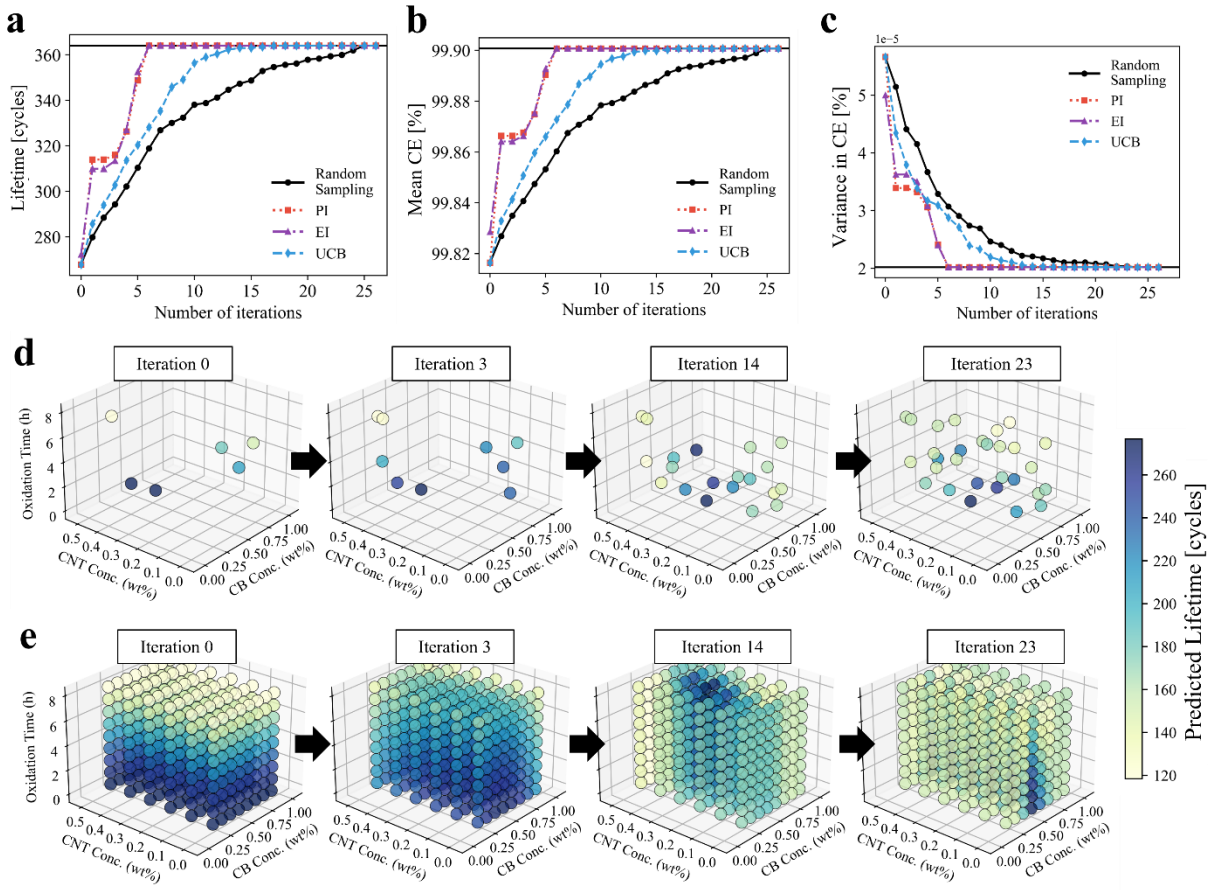

Figure 3: Overview of optimization simulation study results. (a) The maximum measured lifetime is plotted as a function of the current iteration. (b) The mean CE over the first 100 cycles corresponding to the cell with the maximum lifetime from (a), is plotted as a function of the

iteration number. (c) The variance of CE over the first 100 cycles to the cell with the maximum lifetime from (a), is plotted as a function of the iteration number. (d) The evolution of GP model's lifetime predictions for the tested samples. (e) The evolution of the GP model's lifetime predictions over the entire design space, plotted with fewer data points for better visibility.

Figures 3b and 3c show the same trend as **Fig. 3a**. These two plots were generated based on the results in **Fig. 3a**. where the CE and variance in CE are plotted for the current best cell identified via the lifetime metric. We see the same monotonic trend in these two plots, showing that with increased lifetime CE increase and its variance decreases. In **Figs. 3d and 3e**, we visualize the evolution of the GP regression model's predictions over the iterations. The PI acquisition function was used to generate the results in **Fig. 3**. The visualization in **Fig. 3d** provides insight into how the PI acquisition function decided which design variable combinations to test next. The PI acquisition function prioritized testing design variable combinations from iterations three to fourteen with lower oxidation time. The PI acquisition function intelligently identified the overall trend that cells with lower oxidation time performed better. This effect is more noticeable in **Fig. 3e** where the color gradient shows how the GP predicts that cells with lower oxidation time will have longer lifetimes. The algorithm's intelligent understanding of the design space led it to reach the optimum design variable combination roughly five times quicker than the random sampling approach and roughly two times quicker than the UCB acquisition function.

4.2.2. Algorithm sensitivity to initial sample selection

A key factor that influences the performance of the optimization algorithm is the number of initial samples used to fit the GP regression model. To investigate this, we varied the number of randomly selected starting points and repeatedly ran the sequential optimization algorithm to completion. The results, shown in **Fig. 4**, indicate the PI and EI acquisition functions are significantly better suited to our materials optimization problem than the UCB acquisition function is. The PI and EI acquisition functions show minimal change in performance with the number of initial starting points. While the PI and EI acquisition functions certainly perform better, we believe the small size of our dataset influenced this result. Using only 32 total design points is likely insufficient to properly study the initial samples' impact on the algorithm's performance. In the future, it would prove worthwhile to investigate this further. On the other hand, the UCB and random sampling approach follow the same trend. Increasing the initial starting points decreases the iterations required to achieve the optimal nucleation layer design. This result is more consistent with what we expected to observe for the PI and EI acquisition functions. In any case, increasing the number of initial samples should never negatively impact the optimization algorithm. Still, balancing the time required to build and test the samples upfront is essential.

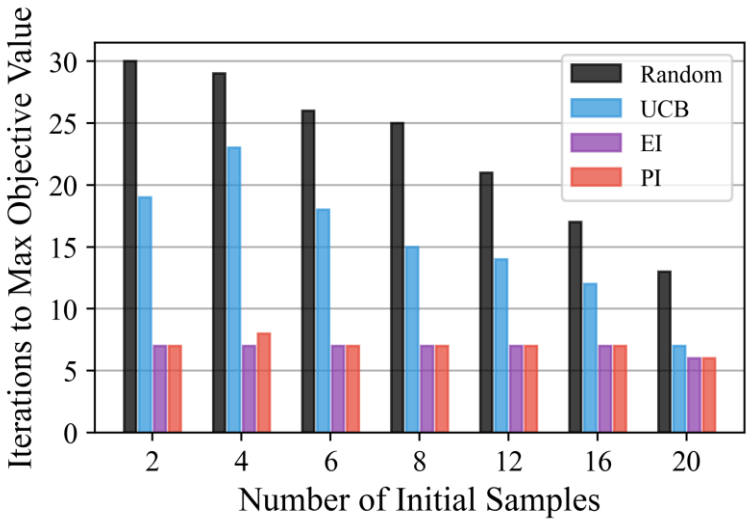

Figure 4: The sequential optimization algorithm's sensitivity to changes in the number of initial samples available. Sequential optimization simulations are repeatedly run using each of the three acquisition functions. Average results from 100 runs are presented.

4.2.3. Acquisition function hyperparameter study

Another key factor that influences the performance of the optimization algorithm is the hyperparameters in the acquisition functions. The simulation studies we presented above were performed using near-optimal values of hyperparameters. Here, we discuss how the hyperparameter values were selected and provide insights on how one might choose suitable parameter values before initiating a battery materials optimization campaign.

We analyzed the algorithm's sensitivity to changes in the hyperparameters by repeatedly running the algorithm to completion over a range of values. We selected the values of the hyperparameters, which were found to reach the maximum objective value in the fewest number of iterations. The PI and EI acquisition functions have a single hyperparameter to tune, namely δ. The δ parameter controls the threshold of improvement, which filters out design points that don't yield an improvement in lifetime greater than the current best cell plus δ. The PI and EI sensitivity analysis results are shown in **Figs. 5a and 5b**. These two acquisition functions perform best with δ greater than ~20, evident by the noticeable decrease in the average number of iterations needed to achieve the maximum objective value for δ values greater than 20. Surprisingly, the EI acquisition function responded to changes in δ in the same way the PI acquisition function did. This result further reinforces the notion that the probability of improvement measured in the PI acquisition function formulation is more important than the magnitude of improvement term in the EI formulation.

The UCB acquisition function has two hyperparameters that need to be optimized, namely β_0 and ϵ. The β_0 parameter controls the initial magnitude of the threshold term, and the ϵ parameter controls the rate of decay over each iteration. Together, these parameters define the threshold as a function of the predicted variance. The results of the UCB hyperparameter sensitivity analysis are shown in **Fig. 5c**. The trend indicates smaller values of β_0 and ϵ perform better. Interestingly, the

hyperparameter values have a noticeable impact on the performance of the UCB acquisition function, even though its best performance is still worse than that of PI and EI.

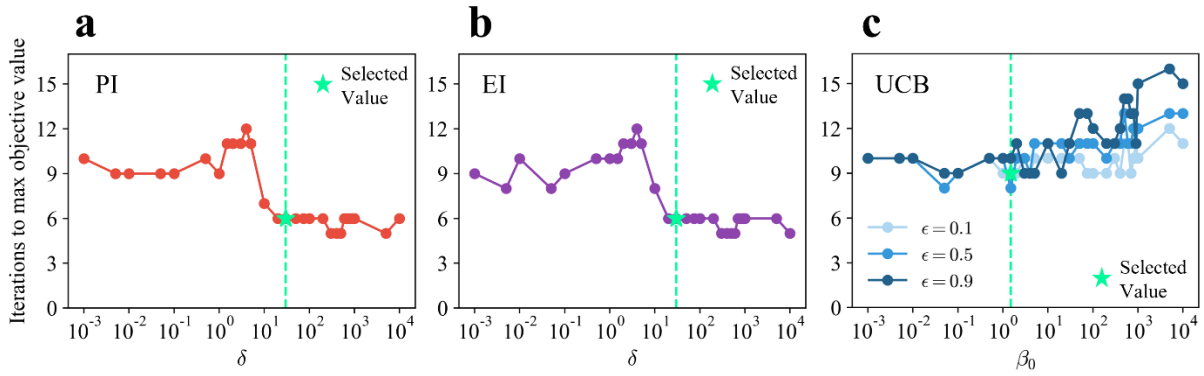

Figure 5: Acquisition function hyperparameter sensitivity analysis results. a) the optimal PI hyperparameter value was selected to be $\delta = 30$, b) the optimal EI hyperparameter value was also selected to be $\delta = 30$, and c) the optimal initial values for the UCB acquisition function were selected to be $\epsilon = 0.1$ and $\beta_0 = 1.5$ where the value of β_k decreases each iteration following the relationship: $\beta_k = \beta_0 \epsilon^k$.

Finding the optimal values for the acquisition function hyperparameters via simulation helps us understand the algorithm's true potential for rapid materials design optimization. However, preemptively selecting the best values for the hyperparameters using post-experiment simulation is infeasible in practice. To get around this, researchers have proposed using simple physics-based and mechanistic models to approximate the objective function for a given set of design variables. Then, simulation experiments, like the ones we conducted, can be run using the simplified model as the ground truth to understand the range of suitable hyperparameter values better. For example, researchers in [12] used a simplified physics-based model for acquisition function hyperparameter selection prior to initiating the optimization campaign. The simplified physics model approximated heat generation within a Li-ion battery cell as a function of the charging currents (design variables). The heat generation was used to compute the expected degradation in the battery cells, which could then be mapped to the total battery lifetime (objective function).

Unfortunately, this approach only works well when the relationship between the design variables and the objective function is well understood. In our case, there is a disconnect between the material-level nucleation layer properties and the cell-level performance that cannot easily be modeled. For example, optimizing the nucleation layer for the maximum surface area does not guarantee long lifetime when tested in a cell. Scenarios like these, where the underlying physics of new materials is not yet understood, are becoming more common. Therefore, we find it essential to share our post-experiment hyperparameter selection results so that others in the community looking to implement optimization algorithms might use our results to inform their algorithm development.

Further investigating the optimal hyperparameter value for the PI acquisition function reveals that it may be related to the cell-to-cell variance. **Figure 6a** shows the distribution of inter-cell

variance, calculated as the standard deviation of the lifetimes for cells of the same nucleation layer design. The mean variation in lifetime between cells of the same design is roughly 25 cycles, which is close to the optimal PI hyperparameter value of $\delta = 30$ (**Fig. 6b**). This suggests that the optimal hyperparameter value is close to the inter-cell variance. In other words, the algorithm needs to be sure a new nucleation layer design will produce a cell with a lifetime greater than the current best cell, even in the presence of unavoidable cell-to-cell variation. We find this result intriguing and believe it would be worthwhile to further investigate the relationship between acquisition function hyperparameters and battery performance variability in future work. Understanding and quantifying the intrinsic cell-to-cell variability that arises when testing cells at identical conditions could help design new customized acquisition functions more suitable for optimizing battery materials. It is likely that optimization speeds could be further increased using custom-designed acquisition functions [48] or by adaptively selecting the best point among the candidates suggested by multiple different acquisition functions [49], [50].

Last, we investigate a scenario where the hyperparameter values cannot be estimated by any means. **Figure 6c** compares each acquisition function's optimization speed using optimized and randomized hyperparameter values. The experiment was conducted with six initial samples and was repeated 100 times. The random sampling approach to design selection is used as a baseline for comparison. The PI and EI acquisition functions show a significant reduction in the run-to-run variance when using a set of optimized acquisition function hyperparameters. On the other hand, the UCB acquisition function shows no significant change in performance from using optimized hyperparameters. This is likely due to our dataset's small size, which inherently makes sensitivity analysis more difficult. Furthermore, tuning the two interacting parameters in the UCB acquisition function is more difficult than tuning the single parameter in PI and EI, adding to the challenge. Using a more rigorous hyperparameter selection method, we may have been able to improve the results of the UCB acquisition function. However, the potential performance gain is small and likely not worth the effort.

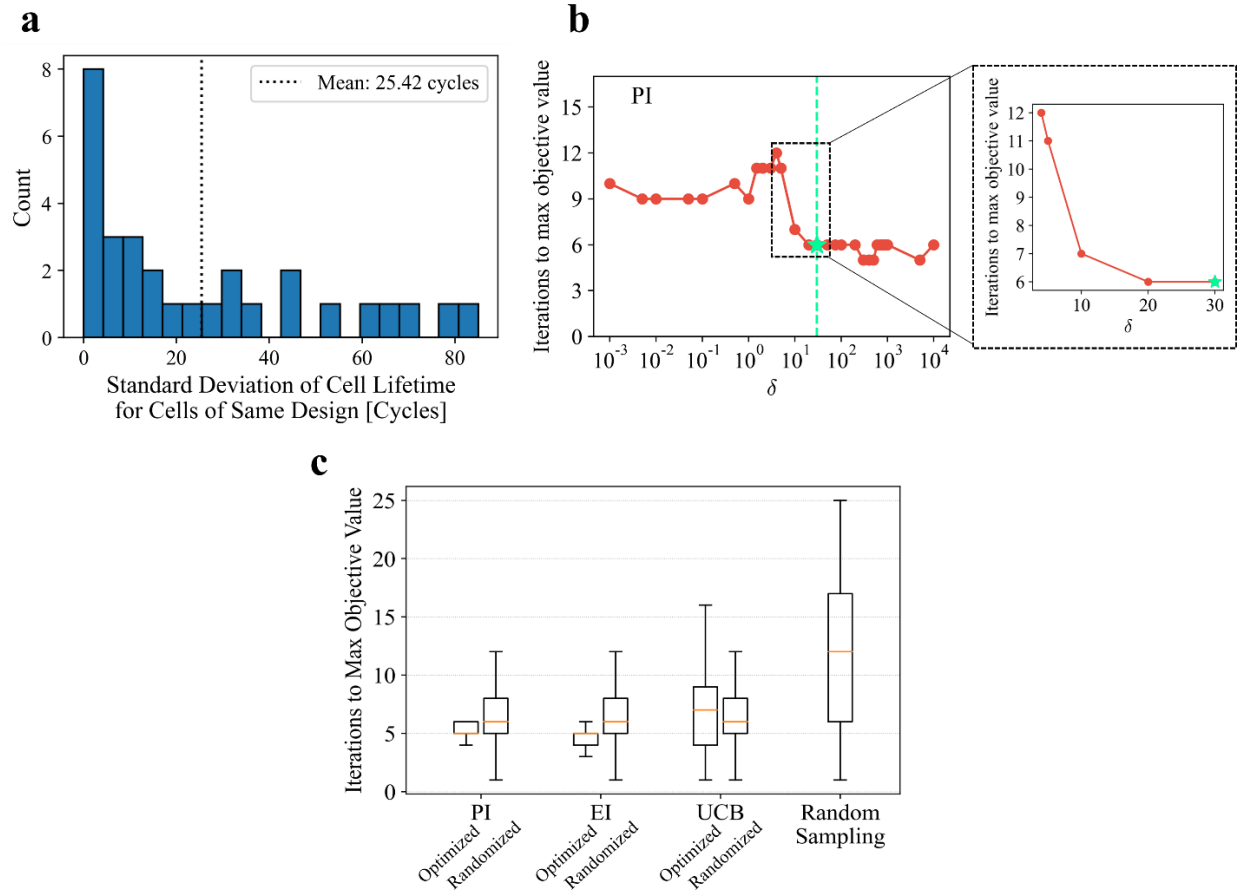

Figure 6: (a) Histogram of cell lifetime variability for cells of the same nucleation layer design. The mean lifetime standard deviation is 25.42 cycles. (b) PI acquisition function hyperparameter sensitivity analysis results from **Fig. 5**, but with an added zoom window showing the trend more clearly. (c) Comparison between optimized and randomized acquisition function hyperparameter values for each of the three acquisition functions, PI, EI, and UCB. Simulations were repeatedly run one-hundred times using six initial samples. The random sampling approach to design selection is used as a baseline for comparison.

Most notably, we observe that even the acquisition functions with randomized hyperparameters performed better than the naïve random sampling approach. This result is great news for others looking to implement optimization algorithms in practice, as a significant performance boost can likely be achieved without performing hyperparameter selection. The main trade-off, however, is that the results are widely varying. The performance increase over the random sampling approach could be as little as two times faster or as much as five times faster. So in any case, it is always better to try and preemptively simulate the optimization algorithm and select appropriate hyperparameter values in advance.

4.3. Experimental Insight for Optimized Performance

In addition to the experimental dataset used thus far to demonstrate optimizing cells' nucleation layer designs based on three parameters: carbon black concentration, CNT concentration, and carbon oxidation, two additional sets of sodium metal coin cells were produced to determine and qualify the choice and selection of the three isolated optimization parameters for use in this study. Details of these additional experimental tests are provided in the **Supplementary Material**. Altogether the experimental results led to a total of 177 cells with 69 different nucleation layer designs. Our efforts in this section focus on using this significant body of results to establish broader correlations that can guide us toward a better understanding of the physical mechanistic behavior of a good nucleation layer.

To uncover the physical mechanisms responsible for cell performance, we turn to the dataset of 177 cells and examine correlations between voltage features visible in standard galvanostatic cycling data and cells' lifetimes. We quantitatively evaluate the Pearson correlation coefficient ρ between different voltage features measured from the half-cell cycling data and a cell's measured lifetime, where lifetime is a measure of cumulative loss of sodium metal based on the CE% over the duration of cycling until a cutoff of 50% sodium loss is reached. Hence, the best-performing cells with the highest CE% will lead to the longest measured lifetime. A high correlation with a low P-value ($p \leq 0.05$) means a statistically significant linear relationship between two variables. For example, a positive correlation indicates that cell lifetime tends to increase as the measured voltage feature increases. We identified six features from the galvanostatic charge-discharge curves that exhibit high correlations with measured lifetime. The correlation analysis results are presented in **Fig. 7.** The two features with the highest Pearson correlation coefficient pertain to the nucleation peak that occurs at the onset of electroplating. They are the time to nucleation, which is the time required to reach the minimum value of the nucleation peak, and the initial slope of the voltage curve, which is the slope of the galvanostatic transient from the onset of electroplating to the minimum point of the nucleation peak. The third best feature is the growth overpotential at which sodium plating occurs. Scatterplots of cell lifetime as a function of the three best features for all 177 cells are shown in **Fig. 7a-c**. The data points are colored from light to dark, representing performance from worst to best. We also overlay regression lines to visualize the trend in the data. The two best feature correlations (see **Fig. 7a-b**) indicate that the best-performing nucleation layer designs exhibit steep slopes that lead to the nucleation peak the fastest.

Other notable features include the nucleation and growth overpotentials, the hysteresis between the electroplating and stripping profiles, and the voltage measured at the very end of the sodium plating process. **Figure 7d** summarizes the correlation coefficient for each feature in a tornado plot, where features with higher correlation coefficients are more strongly related to cell performance. **Figure 7e** visualizes all the features described in **Fig. 7a-d** labeled onto a slip plot, connecting the galvanostatic plating and stripping curves to visualize the sodium loss. In **Fig. 7e**, we refer to the cells using the abbreviations CB, CNT, and Ox, which denote the CB concentration, the CNT concentration, and the oxidation time, respectively.

At first glance of **Fig. 7e**, the differences between the best and worst cells in our optimization study are subtle. However, these two conditions provide an ideal comparison for mechanistic insight into the origin of the two primary parameters correlated to lifetime: time to nucleation peak and initial slope. The main difference between the cells is due to the significant

difference in CE% visualized at the end of the plating profiles in **Fig. 7e**. Additional differences are due to the less prominent features (growth and nucleation overpotential, hysteresis, and end discharge voltage). To better understand the role of these features, we performed additional EIS analysis on cells prepared with the best and worst performing nucleation layers, as this technique provides insight into interfaces and diffusion phenomena in the nucleation layer and how they evolve.

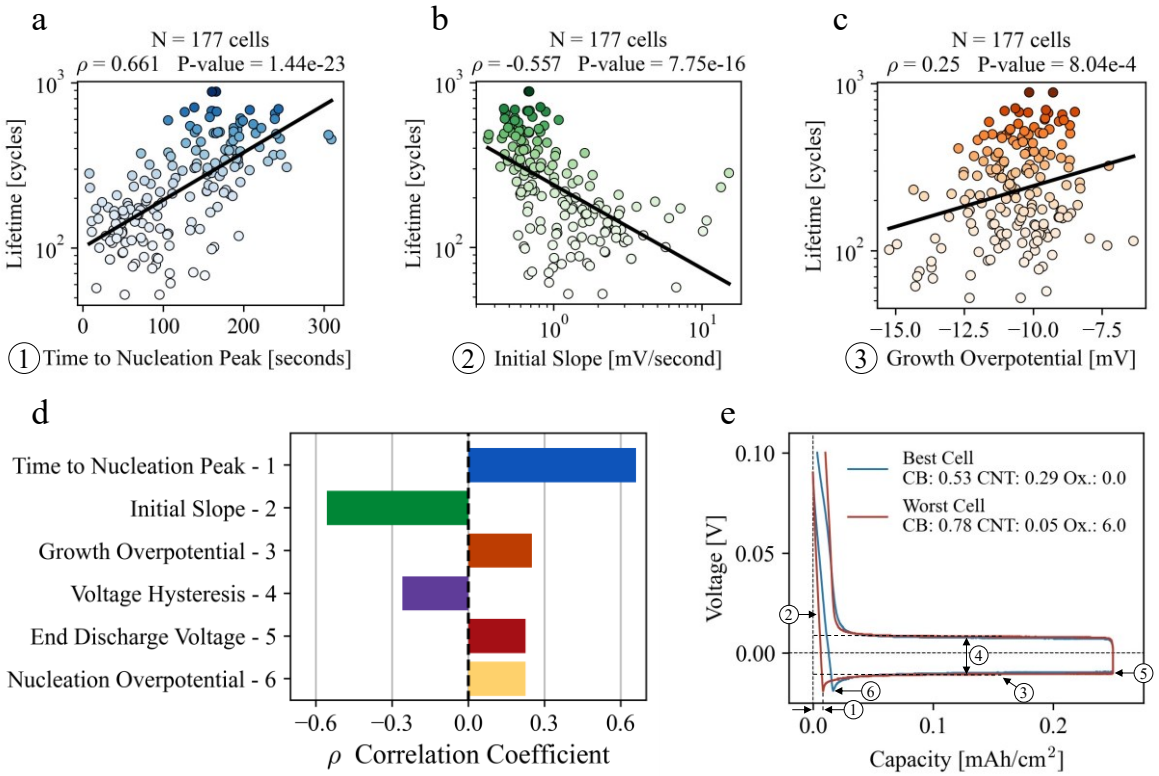

Figure 7: Overview of the most significant correlations from examining the cycle aging data from 177 sodium metal half-cells with 69 nucleation layer designs. The data points are colored from light to dark, representing performance from worst to best. Features with higher correlation coefficients are more strongly related to cell performance. (a) cell lifetime plotted as a function the time to the nucleation peak, with a correlation of 0.661. (b) cell lifetime as a function of the initial slope of the voltage curve, with a correlation of -0.557. (c) cell lifetime as a function of the growth overpotential voltage, with a correlation of 0.250. (d) a tornado plot summarizing the most significant feature-lifetime correlations observed from the 177 cells. (e) the features in (d) are visualized on a slip plot showing the voltage profiles of the best and worst cells from the optimization dataset. All features are calculated as averages over the first 50 cycles.

Figures 8a and **8c** plot the EIS spectra for the best and worst half-cell nucleation layers from the optimization study. Initial EIS data is collected from the cell before metal plating, and subsequent data is measured while the cell is in the charged (stripped) state over the course of ten cycles. We calculate the charge transfer resistance (R_{ct}) by fitting an equivalent circuit to the EIS data and use this information to evaluate the surface kinetics. For the best-performing cell, we

observe during the course of ten cycles, the width of the semicircle, which generally represents the R_{ct} value, gradually decreases to approximately 6.5 Ω. In contrast, the worst-performing cell exhibits a semicircle width of about 9.2 Ω after the first cycle, which remains invariant in later cycles. From these EIS results, the best-performing nucleation layer involves an evolution of the electrode-electrolyte charge-transfer interface over the duration of 10 cycles, whereas the worst-performing nucleation layer does not decrease beyond the 1st cycle. This indicates that interfacial changes occur within the first few cycles, which impact the cell's lifetime and lead to a lower overall resistance. This is analogous to the formation cycles of traditional host anodes such as graphite in Li-ion batteries, where the SEI can continue to improve in the initial few cycles in order to form a robust interphase [51]. **Figure A4** plots the variance of the CE as a function of cycle number. The variance significantly decreases after 10 cycles, suggesting the interface has stabilized. The constant R_{ct} value of the low-performing nucleation layer indicates a stable SEI film is formed in this electrolyte system in accordance with previous studies [21]. Although the SEI film is stable, a further evolution of the interface is associated with improved cell lifetime. **Figure A5** in the Supplementary Material shows an additional high-performing cell with a decreasing R_{ct} value over 10 cycles, further supporting this trend.

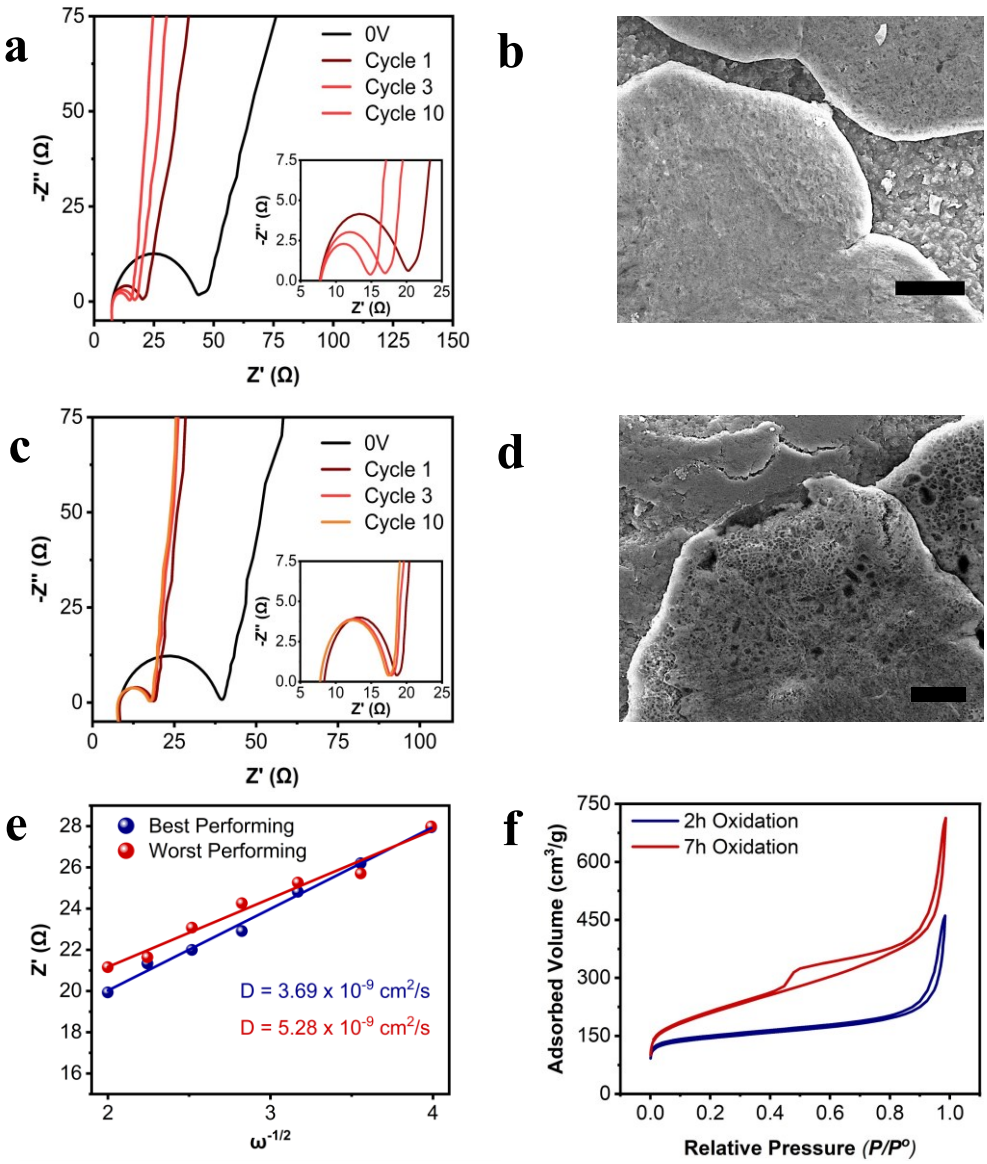

Figure 8: Electrochemical testing and characterization of a high-performing (CB: 0.53, CNT: 0.16, Ox: 2h) and low-performing (CB:0.65, CNT: 0.15, Ox: 7h) nucleation layer design. (a) a Nyquist plots from EIS measurements on the high-performance cell after discharging to 0V and through 10 cycles. The inset figure shows cycles 1-10 in more detail. (b) SEM image of a smooth, faceted sodium deposit on the high-performance substrate (scalebar = 10 μm). (c) Nyquist plots from the low-performance cell. The inset figure shows cycles 1-10 in more detail. (d) SEM image of a rough, porous deposit from the low-performance cell (scalebar = 10 μm). (e) Plots of the real part of impedance (Z′) as a function of the inverse square root of angular frequency ($\omega^{-1/2}$) from the Warburg region of both cells at cycle 10. (f) Nitrogen physisorption isotherms for CB powders oxidized for 2 and 7 hours.

To better understand this relationship, we carried out SEM imaging of the sodium metal deposits from the best and worst-performing nucleation layers (**Figs. 8c-d**). As expected, the planar deposits of sodium metal indicate that nucleation and growth occur in these samples in the kinetic regime [52]. However, evident differences are visible in the sodium deposits' porosity between these two different cases. In the worst-performing nucleation layer, micro-scale pores are distributed throughout the sodium deposits. In contrast, the best-performing nucleation layers result in smoother sodium deposits that visibly exhibit higher volumetric density. Notably, the porous nature of the sodium metal in our poor-performing cell is similar to that discussed elsewhere in the literature for unstable sodium deposition conditions, such as high current density [53]. Two possible reasons for this porous morphology could be (1) diffusion limitations of sodium metal at the nucleation layer-electrolyte interface, or (2) the formation of large dead sodium deposits on the nucleation layer in the initial cycles that physically obstruct localized regions from nucleation and growth of sodium metal.

To evaluate this, we first analyzed the role of diffusion since faster sodium diffusion along the nucleation layer could lead to more uniform metal deposition [34], [54]. We calculated diffusion coefficients using Huggins' method [55] based on the slope of the impedance's real or imaginary parts vs. the frequency's inverse square root, $\omega^{-1/2}$. These plots are shown in **Fig. 8e**, and the diffusion coefficient values for the high-performance and low-performance cells are 3.69×10^{-9} and 5.28×10^{-9} cm^2/s, respectively. Whereas we believe such differences in diffusion coefficient could be related to differences in the surface area of the carbons used in the nucleation layers, as shown in **Fig. 8f**, the measured diffusion coefficient of the worst-performing cell is *higher* than that of the best-performing cell. This elucidates that diffusion limitations are likely not associated with the performance differences between these two cells. Conversely, we believe these results can be explained through dead sodium or SEI that is formed in the first cycle on a poor nucleation layer.

Based on **Fig. 7**, broad correlations across all of our devices with $N = 177$ indicate that a steeper spike into the nucleation peak and less time to nucleation leads to faster sodium loss and degradation. This feature appears from our experiments to be the most important toward determining the effectiveness of a stable nucleation layer. EIS and imaging associate this to a more resistive interface for charge-transfer reactions and correlated increase in porosity of the sodium metal deposits not explained by diffusion-related effects in the two different nucleation layers (**Fig. 8**). In this regard, we hypothesize that in circumstances where the two separate events of SEI formation on the nucleation layer and the initial nucleation of sodium metal are not temporally separated, the formation of significant patches of dead sodium layers can result and impede the stable, dense growth of sodium metal deposits. This could be associated with sodium nuclei formation that fractures off from the nucleation layer due to interfacial stress related to concomitant SEI formation on the nucleation layer. Such dead sodium deposits would produce a physical barrier toward the layer-by-layer growth of dense sodium metal that would explain the increased porosity and poorer long-term cycling behavior of our worst nucleation layers. Overall, we emphasize that this approach of using optimization methods to engineer complex systems such as sodium nucleation layers to obtain high performance, followed by using these systems to evaluate statistically validated design features and mechanisms represents a powerful tool for future battery design efforts.

5. Conclusion

Bayesian optimization and similar machine learning-assisted design optimization strategies are promising solutions to accelerate the development of next-generation batteries. In this work, we demonstrated the ability to quickly optimize three related design variables of a carbon nucleation layer in a sodium metal half-cell. Through repeated simulation, we comprehensively evaluated the performance of the optimization framework and found that using the traditional PI and EI acquisition functions we could reach the optimal nucleation layer design five times faster than a random sampling approach. Additionally, when the acquisition function hyperparameters cannot be optimized in advance, all three of the Bayesian optimization algorithms tested still reached the optimal nucleation layer design roughly two times faster than a random sampling approach.

While the Bayesian optimization algorithm used in this study shows much promise for accelerating battery materials development, we highlight some challenges of applying the algorithm in electrochemical environments. Firstly, the algorithm's speed and efficiency decrease as cell-to-cell variability increases. Cell-to-cell variability typically arises from poor production control or inherently unstable system designs. However, we believe this could be solved by developing new acquisition functions more robust to noise or by testing more cells under the same conditions. Second, the algorithm does not work well for optimizing many ($n >> 20$) design variables, as it requires significantly more experiments to thoroughly cover the entire design space and collect enough data for the model to properly fit the dataset. Last, the algorithm struggles to correctly model design variables that are not continuously variable, as it produces large discontinuities in the model predictions which can cause the algorithm to get stuck during optimization and only focus on a smaller region of the entire design space. Design variables like oxidation time in air and material concentrations work better with the algorithm because they are continuously variable over the entire design space.

Nonetheless, we showed that optimizing the cell-level performance by tuning the material-level properties of a sodium nucleation layer is an effective approach to materials design. We further investigated the relationship between nucleation layer performance and material properties using traditional battery cell characterization methods to analyze the full set of $N = 177$ cells. We find that two primary features, the slope of the nucleation peak and time until the nucleation peak minima in the galvanostatic profile, are highly correlated to cell performance, and indicate these regimes are may be responsible for the majority of dead sodium formation. EIS and imaging characterization correlate poorly performing samples to higher interfacial charge transfer resistance and greater sodium metal porosity that can be attributed to dead sodium that forms due to near-simultaneous SEI formation and sodium metal nucleation on the nucleation layer. The strategies presented in this work broadly apply to battery materials and component optimization and will prove crucial in accelerating the development of next-generation batteries.

Acknowledgments

This work was partly supported by the US National Science Foundation (Grant No. ECCS-2015710) and the Iowa Economic Development Authority through the Iowa Energy Center Grant

Program (Grant No. 21-IEC-007). Any opinions, findings, or conclusions in this paper are those of the authors and do not necessarily reflect the sponsors' views.

References

[1] A. Zeng *et al.*, "Battery technology and recycling alone will not save the electric mobility transition from future cobalt shortages," *Nat. Commun.*, vol. 13, no. 1, p. 1341, 2022, doi: 10.1038/s41467-022-29022-z.

[2] A. Agrawal and A. Choudhary, "Perspective: Materials informatics and big data: Realization of the 'fourth paradigm' of science in materials science," *Apl Mater.*, vol. 4, no. 5, p. 53208, 2016.

[3] K. Rajan, "Materials informatics," *Mater. Today*, vol. 8, no. 10, pp. 38–45, 2005.

[4] Y. Zhong *et al.*, "Mechanistic Insights into Fast Charging and Discharging of the Sodium Metal Battery Anode: A Comparison with Lithium," *J. Am. Chem. Soc.*, vol. 143, no. 34, pp. 13929–13936, Sep. 2021, doi: 10.1021/jacs.1c06794.

[5] J. Seok, J.-H. Hyun, A. Jin, J. H. Um, H. D. Abruña, and S.-H. Yu, "Visualization of Sodium Metal Anodes via Operando X-Ray and Optical Microscopy: Controlling the Morphological Evolution of Sodium Metal Plating," *ACS Appl. Mater. Interfaces*, vol. 14, no. 8, pp. 10438–10446, Mar. 2022, doi: 10.1021/acsami.1c24673.

[6] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, "Machine learning for molecular and materials science," *Nature*, vol. 559, no. 7715, pp. 547–555, 2018, doi: 10.1038/s41586-018-0337-2.

[7] A. Dave *et al.*, "Autonomous Discovery of Battery Electrolytes with Robotic Experimentation and Machine Learning," *Cell Reports Phys. Sci.*, vol. 1, no. 12, 2020, doi: 10.1016/j.xcrp.2020.100264.

[8] S. Sarkar, A. Chaupatnaik, S. D. Ramarao, U. Subbarao, P. Barpanda, and S. C. Peter, "Operando Sodiation Mechanistic Study of a New Antimony-Based Intermetallic CoSb as a High-Performance Sodium-Ion Battery Anode," *J. Phys. Chem. C*, vol. 124, no. 29, pp. 15757–15768, Jul. 2020, doi: 10.1021/acs.jpcc.0c03556.

[9] S. G. Robinson and M. S. Sigman, "Integrating Electrochemical and Statistical Analysis Tools for Molecular Design and Mechanistic Understanding," *Acc. Chem. Res.*, vol. 53, no. 2, pp. 289–299, Feb. 2020, doi: 10.1021/acs.accounts.9b00527.

[10] C. B. Wahl, M. Aykol, J. H. Swisher, J. H. Montoya, S. K. Suram, and C. A. Mirkin, "Machine learning-accelerated design and synthesis of polyelemental heterostructures," *Sci. Adv.*, vol. 7, no. 52, pp. 1–10, 2021, doi: 10.1126/sciadv.abj5505.

[11] M. He *et al.*, "Deterministic inverse design of Tamm plasmon thermal emitters with multi-resonant control," *Nat. Mater.*, vol. 20, no. 12, pp. 1663–1669, 2021, doi: 10.1038/s41563-021-01094-0.

[12] P. M. Attia *et al.*, "Closed-loop optimization of fast-charging protocols for batteries with machine learning," *Nature*, vol. 578, no. 7795, pp. 397–402, 2020, doi: 10.1038/s41586-

020-1994-5.

[13] F. Hildenbrand, F. Aupperle, G. Stahl, E. Figgmeier, and D. U. Sauer, “Selection of Electrolyte Additive Quantities for Lithium-Ion Batteries Using Bayesian Optimization,” *Batter. Supercaps*, vol. 5, no. 7, p. e202200038, Jul. 2022, doi: https://doi.org/10.1002/batt.202200038.

[14] H. Valladares *et al.*, “Gaussian process-based prognostics of lithium-ion batteries and design optimization of cathode active materials,” *J. Power Sources*, vol. 528, p. 231026, 2022, doi: https://doi.org/10.1016/j.jpowsour.2022.231026.

[15] A. Dave *et al.*, “Autonomous Discovery of Battery Electrolytes with Robotic Experimentation and Machine Learning,” *Cell Reports Phys. Sci.*, vol. 1, no. 12, p. 100264, 2020, doi: https://doi.org/10.1016/j.xcrp.2020.100264.

[16] J. S. Kim, D. C. Lee, J. J. Lee, and C. W. Kim, “Optimization for maximum specific energy density of a lithium-ion battery using progressive quadratic response surface method and design of experiments,” *Sci. Rep.*, vol. 10, no. 1, pp. 1–11, 2020, doi: 10.1038/s41598-020-72442-4.

[17] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas, “Taking the human out of the loop: A review of Bayesian optimization,” *Proc. IEEE*, vol. 104, no. 1, pp. 148–175, 2016, doi: 10.1109/JPROC.2015.2494218.

[18] M. Sadoughi, C. Hu, C. A. MacKenzie, A. T. Eshghi, and S. Lee, “Sequential exploration-exploitation with dynamic trade-off for efficient reliability analysis of complex engineered systems,” *Struct. Multidiscip. Optim.*, vol. 57, no. 1, pp. 235–250, 2018, doi: 10.1007/s00158-017-1748-7.

[19] A. Dave, J. Mitchell, S. Burke, H. Lin, J. Whitacre, and V. Viswanathan, “Autonomous optimization of non-aqueous Li-ion battery electrolytes via robotic experimentation and machine learning coupling,” *Nat. Commun.*, vol. 13, no. 1, pp. 1–9, 2022, doi: 10.1038/s41467-022-32938-1.

[20] B. Lee, E. Paek, D. Mitlin, and S. W. Lee, “Sodium Metal Anodes: Emerging Solutions to Dendrite Growth,” *Chem. Rev.*, vol. 119, no. 8, pp. 5416–5460, Apr. 2019, doi: 10.1021/acs.chemrev.8b00642.

[21] Z. W. Seh, J. Sun, Y. Sun, and Y. Cui, “A Highly Reversible Room-Temperature Sodium Metal Anode,” *ACS Cent. Sci.*, vol. 1, no. 8, pp. 449–455, Nov. 2015, doi: 10.1021/acscentsci.5b00328.

[22] L. Ma *et al.*, “Dendrite-free lithium metal and sodium metal batteries,” *Energy Storage Mater.*, vol. 27, pp. 522–554, 2020, doi: https://doi.org/10.1016/j.ensm.2019.12.014.

[23] Y. Deng *et al.*, “On the Reversibility and Fragility of Sodium Metal Electrodes,” *Adv. Energy Mater.*, vol. 9, no. 39, p. 1901651, Oct. 2019, doi: https://doi.org/10.1002/aenm.201901651.

[24] Z. Zhang *et al.*, "Homogenous sdiophilic MoS2/nitrogen-doped carbon nanofibers to stabilize sodium deposition for sodium metal batteries," *Energy Storage Mater.*, vol. 53, pp. 363–370, 2022, doi: https://doi.org/10.1016/j.ensm.2022.09.015.

[25] Z. Li, K. Zhu, P. Liu, and L. Jiao, "3D Confinement Strategy for Dendrite-Free Sodium Metal Batteries," *Adv. Energy Mater.*, vol. 12, no. 4, p. 2100359, Jan. 2022, doi: https://doi.org/10.1002/aenm.202100359.

[26] S. Liu, S. Tang, X. Zhang, A. Wang, Q.-H. Yang, and J. Luo, "Porous Al Current Collector for Dendrite-Free Na Metal Anodes," *Nano Lett.*, vol. 17, no. 9, pp. 5862–5868, Sep. 2017, doi: 10.1021/acs.nanolett.7b03185.

[27] Y. Xie *et al.*, "Encapsulating sodium deposition into carbon rhombic dodecahedron guided by sodiophilic sites for dendrite-free Na metal batteries," *Energy Storage Mater.*, vol. 30, pp. 1–8, 2020, doi: https://doi.org/10.1016/j.ensm.2020.05.008.

[28] C. Chu *et al.*, "Recent advanced skeletons in sodium metal anodes," *Energy Environ. Sci.*, vol. 14, no. 8, pp. 4318–4340, 2021, doi: 10.1039/D1EE01341F.

[29] Z. Sun *et al.*, "Guiding Sodium Deposition through a Sodiophobic–Sodiophilic Gradient Interfacial Layer for Highly Stable Sodium Metal Anodes," *ACS Appl. Energy Mater.*, vol. 4, no. 3, pp. 2724–2731, Mar. 2021, doi: 10.1021/acsaem.1c00016.

[30] K. Lee *et al.*, "A 3D Hierarchical Host with Enhanced Sodiophilicity Enabling Anode-Free Sodium-Metal Batteries," *Adv. Mater.*, vol. 34, no. 14, p. 2109767, Apr. 2022, doi: https://doi.org/10.1002/adma.202109767.

[31] A. P. Cohn, N. Muralidharan, R. Carter, K. Share, and C. L. Pint, "Anode-Free Sodium Battery through in Situ Plating of Sodium Metal," *Nano Lett.*, vol. 17, no. 2, pp. 1296–1301, Feb. 2017, doi: 10.1021/acs.nanolett.6b05174.

[32] A. P. Cohn, T. Metke, J. Donohue, N. Muralidharan, K. Share, and C. L. Pint, "Rethinking sodium-ion anodes as nucleation layers for anode-free batteries," *J. Mater. Chem. A*, vol. 6, no. 46, pp. 23875–23884, 2018, doi: 10.1039/C8TA05911J.

[33] H. Wang, Y. Liu, Y. Li, and Y. Cui, "Lithium Metal Anode Materials Design: Interphase and Host," *Electrochem. Energy Rev.*, vol. 2, no. 4, pp. 509–517, 2019, doi: 10.1007/s41918-019-00054-2.

[34] H. Wang, E. Matios, J. Luo, and W. Li, "Combining theories and experiments to understand the sodium nucleation behavior towards safe sodium metal batteries," *Chem. Soc. Rev.*, vol. 49, no. 12, pp. 3783–3805, 2020, doi: 10.1039/D0CS00033G.

[35] X.-Y. Cui *et al.*, "A Carbon Foam with Sodiophilic Surface for Highly Reversible, Ultra-Long Cycle Sodium Metal Anode," *Adv. Sci.*, vol. 8, no. 2, p. 2003178, Jan. 2021, doi: https://doi.org/10.1002/advs.202003178.

[36] X.-M. Liu *et al.*, "Carbon nanotube (CNT)-based composites as electrode material for

rechargeable Li-ion batteries: A review," *Compos. Sci. Technol.*, vol. 72, no. 2, pp. 121–144, 2012, doi: https://doi.org/10.1016/j.compscitech.2011.11.019.

[37] J. Song, Z. Yu, M. L. Gordin, X. Li, H. Peng, and D. Wang, "Advanced Sodium Ion Battery Anode Constructed via Chemical Bonding between Phosphorus, Carbon Nanotube, and Cross-Linked Polymer Binder," *ACS Nano*, vol. 9, no. 12, pp. 11933–11941, Dec. 2015, doi: 10.1021/acsnano.5b04474.

[38] J.-H. Lee *et al.*, "Strategic dispersion of carbon black and its application to ink-jet-printed lithium cobalt oxide electrodes for lithium ion batteries," *J. Power Sources*, vol. 196, no. 15, pp. 6449–6455, 2011, doi: https://doi.org/10.1016/j.jpowsour.2011.03.041.

[39] C. Fan *et al.*, "Mesopore-dominated hollow carbon nanoparticles prepared by simple air oxidation of carbon black for high mass loading supercapacitors," *Carbon N. Y.*, vol. 160, pp. 328–334, 2020, doi: https://doi.org/10.1016/j.carbon.2020.01.034.

[40] J. Jeong, J. Chun, W.-G. Lim, W. B. Kim, C. Jo, and J. Lee, "Mesoporous carbon host material for stable lithium metal anode," *Nanoscale*, vol. 12, no. 22, pp. 11818–11824, 2020, doi: 10.1039/D0NR02258F.

[41] N. C. Xiao, M. J. Zuo, and C. Zhou, "A new adaptive sequential sampling method to construct surrogate models for efficient reliability analysis," *Reliab. Eng. Syst. Saf.*, vol. 169, no. 2006, pp. 330–338, 2018, doi: 10.1016/j.ress.2017.09.008.

[42] E. Schulz, M. Speekenbrink, and A. Krause, "A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions," *J. Math. Psychol.*, vol. 85, pp. 1–16, 2018, doi: 10.1016/j.jmp.2018.03.001.

[43] M. Li, S. Shen, V. Barzegar, M. Sadoughi, C. Hu, and S. Laflamme, "Kriging-based reliability analysis considering predictive uncertainty reduction," *Struct. Multidiscip. Optim.*, vol. 63, no. 6, pp. 2721–2737, 2021, doi: 10.1007/s00158-020-02831-w.

[44] R. R. Richardson, M. A. Osborne, and D. A. Howey, "Gaussian process regression for forecasting battery state of health," *J. Power Sources*, vol. 357, pp. 209–219, 2017, doi: 10.1016/j.jpowsour.2017.05.004.

[45] M. Sadoughi, M. Li, and C. Hu, "Multivariate system reliability analysis considering highly nonlinear and dependent safety events," *Reliab. Eng. Syst. Saf.*, vol. 180, pp. 189–200, 2018, doi: https://doi.org/10.1016/j.ress.2018.07.015.

[46] H. J. Kushner, "A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise," 1964.

[47] J. Mockus, "The Bayesian approach to global optimization," in *System Modeling and Optimization*, Springer, 1982, pp. 473–481.

[48] E. Brochu, V. M. Cora, and N. de Freitas, "A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical

Reinforcement Learning," 2010, [Online]. Available: http://arxiv.org/abs/1012.2599.

[49] M. Hoffman, E. Brochu, and N. De Freitas, "Portfolio Allocation for Bayesian Optimization.," in *UAI*, 2011, pp. 327–336.

[50] B. Shahriari, Z. Wang, M. W. Hoffman, A. Bouchard-Côté, and N. de Freitas, "An entropy search portfolio for Bayesian optimization," *arXiv Prepr. arXiv1406.4625*, 2014.

[51] P. G. Kitz, M. J. Lacey, P. Novák, and E. J. Berg, "Operando investigation of the solid electrolyte interphase mechanical and transport properties formed from vinylene carbonate and fluoroethylene carbonate," *J. Power Sources*, vol. 477, p. 228567, 2020, doi: https://doi.org/10.1016/j.jpowsour.2020.228567.

[52] L. Guo, G. Oskam, A. Radisic, P. M. Hoffmann, and P. C. Searson, "Island growth in electrodeposition," *J. Phys. D. Appl. Phys.*, vol. 44, no. 44, p. 443001, 2011, doi: 10.1088/0022-3727/44/44/443001.

[53] P. M. Bayley, N. M. Trease, and C. P. Grey, "Insights into Electrochemical Sodium Metal Deposition as Probed with in Situ 23Na NMR," *J. Am. Chem. Soc.*, vol. 138, no. 6, pp. 1955–1961, Feb. 2016, doi: 10.1021/jacs.5b12423.

[54] J. Yan *et al.*, "3D printed rGO/CNT microlattice aerogel for a dendrite-free sodium metal anode," *J. Mater. Chem. A*, vol. 8, no. 38, pp. 19843–19854, 2020, doi: 10.1039/D0TA05817C.

[55] C. Ho, I. D. Raistrick, and R. A. Huggins, "Application of A-C Techniques to the Study of Lithium Diffusion in Tungsten Trioxide Thin Films," *J. Electrochem. Soc.*, vol. 127, no. 2, pp. 343–350, 1980, doi: 10.1149/1.2129668.